8-2004

Developing Habitat Suitability Models: An Example From Great Smoky Mountains National Park, Tennessee and North Carolina, USA, Using the Land Snail *Vitrinizonites latissimus* Lewis

Andrew Strom Dye
University of Tennessee, Knoxville

To the Graduate Council:

I am submitting herewith a thesis written by Andrew Strom Dye entitled "Developing Habitat Suitability Models: An Example From Great Smoky Mountains National Park, Tennessee and North Carolina, USA, Using the Land Snail *Vitrinizonites latissimus* Lewis." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.

Henri D. Grissino-Mayer, Major Professor

We have read this thesis and recommend its acceptance:

Sally P. Horn, Ken H. Orvis

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting a thesis written by Andrew Strom Dye entitled "Developing Habitat Suitability Models: An Example From Great Smoky Mountains National Park, Tennessee and North Carolina, USA, Using the Land Snail *Vitrinizonites latissimus* Lewis." I have examined the final paper copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.

Dr. Henri D. Grissino-Mayer, Major Professor

We have read this thesis
and recommend its acceptance:

Sally P. Horn

Dr. Sally P. Horn

Dr. Ken H. Orvis

Accepted for the Council:

Vice Chancellor and Dean of
Graduate Studies

DEVELOPING HABITAT SUITABILITY MODELS: AN EXAMPLE FROM GREAT SMOKY MOUNTAINS NATIONAL PARK, TENNESSEE AND NORTH CAROLINA, USA, USING THE LAND SNAIL *VITRINIZONITES LATISSIMUS* LEWIS

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Andrew Strom Dye
August 2004

Dedication

This thesis is dedicated to my family, who has blessed me with a multitude of opportunities throughout my life for which I am eternally grateful. Their love, support, and guidance have led me to where I am today, and this thesis would not be possible without them. They helped instill a love for the Great Smoky Mountains at an early age that has carried over to the pages of this thesis. And I also dedicate this thesis to my late grandfather Robert B. Strom and grandmother Genevieve Strom, who also shared a love of family, faith, and of the Great Smoky Mountains.

Acknowledgements

I would like to thank several people for their assistance and support of this project. First, I would like to thank my advisor, Dr. Henri Grissino-Mayer, for his guidance throughout my time in the Master's program. His direction and advising through this thesis and my studies at the University of Tennessee were invaluable to my education. I would also like to thank my other committee members, Dr. Sally Horn and Dr. Ken Orvis, for their help and suggestions to help strengthen this document. I am grateful for the support and guidance from Mr. Keith Langdon of the National Park Service and Ms. Jeanie Hilton of Discover Life in America. I would like to thank Mr. Dan Dourson and Ms. Stephanie Stocks for sharing their knowledge and expertise about land snails and biological sampling in the Smoky Mountains. I am also thankful for the support and friendship of Dr. John Rehder during my time at UT.

I am very grateful for all the volunteers who helped me collect the data for this thesis and especially Ms. Michelle Prisby of the Great Smoky Mountains Institute at Tremont and Mr. Kurt Butefish of the Tennessee Geographic Alliance for helping coordinate the efforts of these volunteers. I would like to thank the teachers and students of Alliance workshops, attendants of the Tremont workshops, the Wilderness Volunteer group, members of the National Park Service, students of the Upward Bound Program, Daniel Bradford, Hank and Margaret Dye, and fellow UT students Ben Mainor, Amy Hill, Allie Ridenour, Andrew Parrish, Anna Compton, Chris Knutson, Wes Sternberg, Martin Lafrenz, Claire Jamieson, Alan Jolly, and Jody Sumner for their assistance in data collection and other aspects of this thesis research.

Abstract

Great Smoky Mountains National Park (GSMNP) is one of the foremost areas of biological research in the United States. The spatial and temporal aspects of the data collected in biological research provide many opportunities for implementing Geographic Information System (GIS) models. Many biological studies in recent years have sought to describe the relationships between organism distribution and habitat variables.

This thesis analyzes the spatial distribution of the land snail *Vitrinizonites latissimus* Lewis as an example of how GIS-based habitat suitability models can be derived from data collected in GSMNP. Seven habitat variables including slope, aspect, elevation, soils, vegetation, geology, and logging history were analyzed in relation to *V. latissimus* occurrence data to identify the primary components of the snail's habitat. The variables that were found to have an influence on the distribution of the snail (elevation, soils, and vegetation) were input into ArcGIS Spatial Analyst's raster calculator to derive the suitable habitat zones for the organism within GSMNP. Spruce-fir and northern hardwood forests, Breakneck-Pullback, Oconaluftee-Guyot Chiltoskie, and Luftee-Anakeesta soil units, and elevations above 1400 m are the primary habitat characteristics for *V. latissimus.* There are 8,872 ha of optimal habitat conditions located in GSMNP.

The model procedures can be applied to all biological data collected in the park and will help park managers predict the effects of environmental degradation such as acid deposition or invasive species in the park. Habitat models will also provide park managers a way to predict the environmental impacts of developments in the park such as roads and parking lots. Models as described in this thesis can serve as an important tool for conservation biology in GSMNP and throughout the entire National Park system.

Table of Contents

Chapter One: Introduction 1

1.1 Purpose....1
1.2 Justification....2
1.3 Overview of the All Taxa Biodiversity Inventory....7
1.4 Geodatabase Introduction....8
1.5 Scope and Limitations....9
1.6 Research Objectives....10
1.7 Thesis Organization....11

Chapter Two: Literature Review 13

2.1 GIS and Environmental Modeling....13
2.2 Wildlife Biology and GIS....15
2.2.1 Case Study: Adverse Effects of Acid Rain on the Distribution of the Wood Thrush *Hylocichla mustelina* in North America (Hames *et al.* 2002)....16
2.2.2 Case Study: Using a GIS Model to Assess Terrestrial Salamander Response to Alternative Forest Management Plans (Gustafson *et al.* 2001)....18
2.2.3 Case Study: Least-Cost-Path Corridor Analysis Analyzing Wildlife Movement Corridors in Montana Using GIS (Craighead and Walker 1997)....19
2.3 Habitat Suitability Models....21
2.3.1 Case Study: A GIS Analysis and Model of Suitable Gray Wolf Habitat in the Northern Rockies (Brehme 2001)....22
2.3.2 Case Study: A GIS-based Habitat Model for Wood Thrush, *Hylocichla mustelina,* in Great Smoky Mountains National Park (Shriner *et al.* 2002)....24
2.3.3 Case Study: A GIS Model for Identifying Potential Black-Tailed Prairie Dog Habitat in the Northern Great Plains Shortgrass Prairie (Proctor *et al.* 1998)....26
2.3.4 Case Study: Prediction of Thistle-infested Areas in Badlands National Park Using a GIS Model (Price and Tinant 2000)....28
2.4 GPS/GIS Integration for Environmental Research....29
2.4.1 Case Study: Mapping Southern Oregon's Noxious Weeds with GPS (Kiser and Williams 2000)....30
2.4.2 Case Study: Mapping Urban Deer Populations using GPS and GIS (Blake and Brown 2001)....31
2.4.3 Case Study: GIS Mapping Aids Understanding of Elephant Seal Behavior (Galimberti and Sanvito 1999)....32

Chapter Three: Study Area Description 33

3.1 Great Smoky Mountains National Park....33

3.2 Geology of the Great Smoky Mountains....35
3.3 Soils of the Great Smoky Mountains....36
3.4 Climate of the Great Smoky Mountains....39
3.5 Forest Cover within the Great Smoky Mountains....43
3.6 Topographic Characteristics of the Great Smoky Mountains....47
3.7 Disturbance History of the Great Smoky Mountains....49

Chapter Four: Methods and Procedures 55
4.1 Site Selection....55
4.2 Data Acquisition....55
4.3 Data Organization....60
4.3.1 Geodatabase Structure....60
4.3.2 Projection Information....60
4.3.3 Vector Data....62
4.3.4 Grid Data....65
4.4 Data Preparation....66
4.5 Statistical Analysis....68
4.6 Data Processing....71
4.7 Species Abundance Model....72

Chapter Five: Results 73
5.1 Cross-tabulation and Chi-Squared Results....73
5.2 Assessing Variable Weight Assignment Using Logistic Regression....83
5.3 Variable Reclassification Results....85
5.4 Habitat Suitability Model Results....91
5.5 Habitat Suitability Model Verification....91
5.6 Possible Extension of the Model....94

Chapter Six: Discussion 97
6.1 The Habitat Suitability Model....97
6.2 Habitat Variables Not Included in the Model....101
6.3 Importance to Management....103
6.4 Statistical Considerations....110
6.5 Possible Sources of Error....114

Chapter Seven: Conclusions 117
7.1 Preferred Habitat Characteristics of *Vitrinizonites latissimus* in GSMNP...117
7.2 Distribution of *V. latissimus* in GSMNP....119
7.3 The GIS-based Model....121

Literature References 125
Appendices 135
Appendix A: *Vitrinizonites latissimus* Description....137
A.1 Taxonomy....139

A.1 Taxonomy....139
A.2 General Description of the Species....140

Appendix B: Habitat Suitability Model Tutorial....143
B.1 Merging Individual Trail Files into One Master Trail File....145
B.2 Spatially Joining Polygon Data to Points....147
B.3 Joining Raster Data to the Point File....149
B.4 Reclassifying Environmental Factor Grids....151
B.5 Evaluating Habitat Suitability Areas....153

Vita **155**

List of Tables

Table 3.1 Soil units of GSMNP (data obtained from the USGS) 38

Table 3.2 Monthly weather at Gatlinburg, elevation 446 m (Wuerthner 2003)................42

Table 3.3 Monthly weather at Clingman's Dome, elevation 2,024 m (Wuerthner 2003)..42

Table 4.1 Data dictionary fields..59

Table 4.2 GIS data layers used for the project..63

Table 5.1 Reclassification values for geology..74

Table 5.2 Cross-tabulation results for geology...75

Table 5.3 Reclassification values for vegetation..76

Table 5.4 Cross-tabulation results for vegetation...77

Table 5.5 Reclassification values for soils..79

Table 5.6 Cross-tabulation results for soils...80

Table 5.7 Class values for logging history..81

Table 5.8 Cross tabulation results for logging history..82

Table 5.9 T-test results for slope, aspect, and elevation..84

List of Figures

Figure 3.1 Great Smoky Mountains National Park..34

Figure 3.2 Geology of GSMNP (data obtained from the NPS)..37

Figure 3.3 Soil classes of GSMNP (data obtained from USGS)......................................40

Figure 3.4 Vegetation of GSMNP (data obtained from the NPS)....................................44

Figure 3.5 Elevation ranges of GSMNP (data obtained from the NPS)...........................48

Figure 3.6 Logging history of GSMNP (data obtained from the NPS)............................50

Figure 4.1 Systematic sampling grid...56

Figure 4.2 Sample Sites in GSMNP...57

Figure 4.3 Flow chart of GIS data organization..61

Figure 5.1 Reclassified vegetation layer..86

Figure 5.2 Reclassified soil layer...88

Figure 5.3 Reclassified elevation layer..89

Figure 5.4 Histogram of elevation (m) and presence of *V. latissimus*..............................90

Figure 5.5 Habitat suitability model for *V. latissimus*..92

Figure 5.6 Verification map of suitability zones and the presence/absence of *V. latissimus* 93

Figure 5.7 Species abundance map for Clingman's Dome, GSMNP..............................95

Figure 6.1 Scatterplot of elevation (m) and soil units of GSMNP.................................111

Figure 6.2 Scatterplot of elevation (m) and geology classes in GSMNP........................112

Figure 6.3 Scatterplot of elevation (m) and vegetation types in GSMNP.......................113

Figure A.1 Shell of *V. latissimus*...141

Figure A.2 *V. latissimus* in GSMNP..141

List of Plates

Plate One: Habitat Suitability Map for *V. latissimus* in GSMNP............................In Pocket

List of Abbreviations

ATBI: All Taxa Biodiversity Inventory

DEM: Digital Elevation Model

ESRI: Environmental Systems Research Institute

GIS: Geographic Information Systems

GPS: Global Positioning Systems

GSMNP: Great Smoky Mountains National Park

NAD: North American Datum

NPS: National Park Service

NRCS: Natural Resource Conservation Service

PDOP: Position Dilution of Precision

SAMI: Southern Appalachian Mountains Initiative

SAS: Statistical Analysis Software

SDSS: Spatial Decision Support System

SPSS: Statistical Package for Social Sciences

USGS: United States Geological Survey

UTM: Universal Transverse Mercator

χ^2: Chi Square Analysis

Chapter One
Introduction

1.1 Purpose

To properly manage the diverse and abounding life within Great Smoky Mountains National Park (GSMNP), National Park Service conservationists require knowledge of not only the ecology, but also the distribution of particular organisms. Identifying habitat characteristics for a particular species allows researchers to target specific areas of the park for study so they can more efficiently sample the park. A Geographic Information System (GIS) can generate habitat suitability models so that much of the guesswork is eliminated and time in the field can be far more productive. GIS technology can relate information about an organism with the environmental characteristics that define its habitat, and it helps users visualize data in meaningful ways. Many applications exist for applying such habitat suitability models to conservation and biologic inventory efforts in the park and throughout the world.

In this study, I analyzed the relationship between environmental factors and the land snail *Vitrinizonites latissimus* Lewis (Appendix A) within GSMNP. Environmental factors (such as geology, vegetation, soils, slope, aspect, logging history, and elevation) play a major role in land snail distributions. The purpose of this study was to examine selected portions of the park to identify specific habitat characteristics preferred by this land snail. By identifying and classifying the habitat characteristics of this species, I developed a GIS habitat suitability model. The purpose of this habitat model approach is to offer researchers and decision makers a tool that they can integrate within a wide range of applications in GSMNP. This study produced a step-by-step guide to developing a model for most any species found within the park. The guide offers users who are otherwise not familiar with GIS a way to implement this model in their own research.

This study also used GIS to compare species abundance between the modeled species *V. latissimus* and another snail species, *Stenotrema depilatum* Pilsbry. *V. latissimus* has a shell made primarily of protein (Langdon 2001) while *S. depilatum* has a calcium-based shell. The protein-based shell requires less calcium for development than the calcium-based shell and is therefore less affected by the acid deposition that threatens the upper elevations of GSMNP. This aspect of the research represents groundwork for future studies of the effects of acid deposition in the park on land snail populations.

The development of the habitat suitability model makes the benefits of using GIS for conservation biology evident. A GIS makes it possible to not only build basic maps from vast datasets, but also helps the analysis of species richness and population interactions, and can be a primary tool to prioritize areas for conservation based on habitat niches (Peterson *et al.* 2000). GIS software provides fast, accurate analysis for a wide range of data. Models such as the one detailed in this study can eventually become an integral part of wildlife conservation as they are dynamic in nature and offer efficient methods.

1.2 Justification

When resource managers study a species, they often know the general characteristics of the habitat in which the species is found, but not where these habitats are located for a given region. Research endeavors such as the All Taxa Biodiversity Inventory (ATBI) currently being conducted in GSMNP utilize the knowledge and skills of expert biologists all over the world. The ATBI seeks to classify every living organism in the park as well as gain species distribution information about each organism. These experts know the biology of their particular taxa, but often times know little about the geographic distributions of the specific taxa in the park and the spatial relationships that are present with certain environmental variables. This model will serve as a guide to help researchers identify suitable habitats in which to perform their studies as well as picture their data in meaningful ways. The model approach will also allow researchers to more efficiently sample the park, as they will be able to target specific areas where their species are likely to be found. Habitat suitability models can be produced from the vast amounts of spatial, biologic data obtained through the ATBI so that decision makers can clearly see solutions to questions and issues that arise pertaining to conservation in the park. Knowledge of how environmental variables define an organism's habitat makes conservation of that organism easier.

The primary focus of this work is to facilitate conservation and research efforts in GSMNP, but the model created has wider potential. Conservation managers can model possible changes in the park with the GIS, based on which characteristics are important to species' distributions, and they can assess how these changes will affect the organism. For example, eastern hemlock (*Tsuga canadensis* (L.) Carr.) stands may decline because of the invasive hemlock wooly adelgid (*Adelges tsugae* Annand). A GIS model similar to that designed for this study can determine how the decline will affect certain organisms that rely on this tree species as a vital part of their environment. The ability to predict what will happen if certain characteristics that are important to the biodiversity of the park are lost gives managers an advantage. The habitat suitability models will help organism reintroduction efforts as well. Managers will be able to see where suitable habitats are located and be able to tell how the reintroduction of a species will affect other organisms already occupying that space.

Such habitat models will help predict species migration by integrating habitat models that potentially have an influence on one another. For example, elk (*Cervus elaphus* Erxleben) were reintroduced in the Cataloochee area of GSMNP in the winter of 1998. Managers can predict where the elk populations might disperse by looking at the distribution of Turk's Cap's Lily (*Lilium superbum* L.), a plant species heavily preferred by elk. These models will also let managers predict how altering the landscape will affect species in the park. If park planners decide a new road is needed in the park, its effects on the biology of the park can be modeled with the integration of these habitat suitability models.

Snails are a vital part of the GSMNP diversity as they are a major food source and potential calcium source for organisms such as birds and small mammals. A decline in the snail populations affects organisms in higher trophic levels that depend on the snails as a source of calcium (Hotopp 2002). Snail shells are often found in forest passerines' nests as the main calcium-rich food source (Graveland 1996). An understanding of why these snails are found where they are is important to help predict where to find other species, *i.e.* black-capped chickadees (*Parus atricapillus* L.), wood thrush (*Hylocichla* *mustelina* Gmelin), and red-eyed vireo (*Vireo olivaceous* L.), that are harder to sample. Snail distributions play a key role in understanding ecology in the mountains. Surveying and identifying life history characteristics of each species within a region as diverse as GSMNP is an overwhelming, time-consuming task, but this GIS model could help resolve this problem.

Land snails are very susceptible to acid deposition, which is common in the higher elevations of GSMNP (Langdon 2001). If soil calcium is important to land snails, then land snail communities may be sensitive to reductions in soil calcium, including those caused by atmospheric deposition. Calcium is one of the cations that is stripped from the environment by acid precipitation (Simons 2001). Sulfur and nitrogen deposition rates within the park are some of the highest found in North America. The pollutants come not only in the form of rainfall, but also from large amounts of dry particles and cloud vapor. Natural rainfall has an average pH of 5.0–5.6; the average annual pH of park rainfall is 4.5. Cloud water pH levels in the upper elevations of GSMNP average 3.5 and have been measured as low as 2.0 (Renfro 2001). Some soils in the higher elevations are undergoing advanced stages of nitrogen saturation, which causes leaching of nutrients like calcium (NPS 2001). Investigating the effects of acid deposition on land snail populations is beyond the scope of this study, but the study will provide a starting point for investigating the interaction. Knowledge concerning the preferred habitat for specific land snail populations will provide baseline information for analyzing associations between acid deposition and geographic distribution.

All snails found at a sample sites were collected during the initial sampling stages, and it was determined that modeling the habitats of all species was beyond the scope of this study. The land snail species *V. latissimus* was chosen for the study due to its easy identification and relative abundance in the park. The fact that the shell of this snail is comprised of protein instead of calcium makes it a unique land snail species to study, especially with the acid deposition levels found in GSMNP. Snails are good indicators of environmental health since they are not very mobile and are affected by alterations in local habitat quality. *V. latissimus* may not be as dependant on soil calcium as other snail species are, which may cause the snail to be found in greater relative abundance in areas of low calcium content such as the upper elevations of GSMNP.

Biological GIS models can help determine habitat characteristics for a wide range of species found in GSMNP. Presence/absence data for an individual species throughout the park, combined with detailed maps of major vegetation types, geologic classes, elevation, slope, aspect, and other habitat characteristics, can help us to understand a species' distribution even if very little is previously known about the species.

Understanding where the abundant wildlife is found in the park helps explain why it is found there; this understanding is a vital part of conservation and management efforts. Park conservationists can determine which environmental characteristics should be given priority for conservation by knowing what threatened or endangered species need in order to survive. Biological habitat models give planners and decision makers a tool to assist them in efforts to optimize biodiversity conservation and preservation (Peterson *et al.* 1999).

1.3 Overview of the All Taxa Biodiversity Inventory

ATBI researchers will be the primary target users of this model. The ATBI is a research effort initiated in 1998 that attempts to identify all organisms living in GSMNP. Information gathered in this project will help park managers make decisions that are critical to protecting the natural resources in the park from a growing number of threats. The goals of the ATBI include:

- Developing a comprehensive "checklist" of life forms in GSMNP;
- Gathering data to create distribution maps for each species;
- Compiling natural history information on each species, including abundance in the park, its role in the greater ecosystem of the park, photographs of each life stage, and recordings of calls or sounds;
- Presenting educational materials to teachers and students, scientists, land managers, and others via the Internet and a variety of media.

Researchers from around the world have come and continue to come to the park to survey taxa within their area of expertise, which helps the Park Service better monitor and conserve the community structure. An estimated 90% of the invertebrate species in the park are still unknown, including 50% of the snail population (Discover Life in America 2003).

This major survey is the largest biologic inventory endeavor ever undertaken, and it serves as a pilot project that will set the standard for conservation efforts around the nation and throughout the world. Scientists have estimated that more than 100,000 plants and animals live in GSMNP, but at the beginning of this project, fewer than 10,000 had been identified (Sharkey 1997). At the time of the seventh annual ATBI conference in Gatlinburg, Tennessee, in December 2003, 410 species new to science had been found in the park and 2,955 species new to the park had been recorded (White 2003). A standard approach to developing habitat suitability models will allow researchers who have a limited amount of time to spend in the park to efficiently survey the park and make the best use of their time. As the amount of data collected by the ATBI continue to increase, researchers are searching for ways to manage and visualize these data. This was a common theme among the progress reports at the 2003 ATBI conference. The scope and aim of this thesis is to answer the call of researchers by offering them a means to represent and examine their data in meaningful ways.

1.4 Geodatabase Introduction

All the data for this study were managed using Environmental Systems Research Institute's (ESRI) ArcCatalog platform. The geodatabase design within ArcCatalog is a hierarchical form of data management that provides a seamless view of geographic data. The geodatabase is the top-level unit of geographic data; it is a collection of various datasets, feature classes, object classes, and relationship classes (Zieler 1999). All vector GIS data are stored as feature classes within feature datasets. Feature datasets are a collection of similar feature classes and are used to ensure data integrity and maintain a common coordinate system. All feature classes in this project are simple feature classes meaning that they are a collection of features with similar geometries without any topological associations defined between them. Object classes, tables that describe geographic features, are also stored within feature datasets. Raster data, such as grid files, are stored outside of the geodatabase in a common folder.

ArcCatalog provides an environment to allow the investigation to explore, assess, manage, and build geographic data. For this project, it helped create and form new data, imported data from outside sources, searched for data, defined the coordinate system, and launched GIS operations. The data could be previewed and edited within ArcCatalog as either spatial geography or table information. It is difficult to keep track of edits made in and out of the GIS when analyzing large amounts of data as were dealt with in this project. With the geodatabase design, this problem was alleviated, and data integrity was ensured. Data do not have to be imported and exported to database programs for editing, so less potential exists for data to be lost in transition.

1.5 Scope and Limitations

This study primarily focuses on the developing habitat suitability models and it features the land snail species *Vitrinizonites latissimus.* The model focuses solely on this species because it was easily identifiable in the field and was found in relative abundance within the park. Not all snails observed at each site were included in the model. The study will compare species abundance of *V. latissimus* and *Stenotrema depilatum* and their association with acid deposition areas. The more abundant of the two was used for the habitat suitability model.

ATBI researchers collect vast amounts of biological data during their surveys, but rarely are able to visually represent their data in meaningful ways. This thesis is meant to develop a model approach that is easy to apply to the biological data unique to GSMNP. The model procedure has been developed in a manner that is easy for researchers not familiar with GIS to follow and will serve as a guide for future researchers to help to optimize their sampling time in the park, to visualize habitat data, and to observe associations with environmental properties for a wide range of organisms.

At the time of model development, detailed park geology and vegetation data were not available. Once these data become available, these GIS models will gain increased accuracy. All data used in this model were the most detailed datasets available at the time. These data may be obtained from the GSMNP GIS staff at the Twin Creeks research facility in Gatlinburg, Tennessee.

1.6 Research Objectives

This research has the following specific research objectives:

1) Determine the major habitat characteristics for *Vitrinizonites latissimus* by querying the GPS sample site data, and overlaying it with habitat characteristic data such as elevation, geology, vegetation, soils, logging history, slope, and aspect to identify which characteristics of these variables are found in association with the presence of the species.
2) Determine the distribution and total area of suitable habitat locations for *V. latissimus* by using statistical analysis that assesses which characteristics will be used in the model.

3) Demonstrate how GIS technology can be a major tool in modeling habitat distributions and species abundance within GSMNP and show that GIS is an efficient and viable option for storing and managing GSMNP data.
4) Produce a step-by-step guide to develop habitat suitability models for any species found in GSMNP so that biological researchers who are unfamiliar with GIS will be able to utilize the technology as a valuable tool for their research.

1.7 Thesis Organization

This thesis is organized into seven chapters. Chapter Two is a review of the literature, and provides discussion on topics involving GIS and environmental research as well as specific case studies that showcase GIS in conservation biology. Chapter Three gives a detailed description of GSMNP and its environmental characteristics. Chapter Four gives a description of the methods and procedures used in this thesis, and includes a description of the data and data structures used in the study. A more detailed, step-by-step tutorial for the habitat suitability model is presented Appendix B. Chapter Five provides the results of the model and statistical analysis conducted on the variables. Chapter Six offers discussion about the results of the model and statistics as well as a discussion about the possible applications of the model in GSMNP conservation efforts. Chapter seven presents the conclusions of this thesis.

Chapter Two
Literature Review

2.1 GIS and Environmental Modeling

GIS technology is a growing field with applications in a variety of disciplines, including transportation, business, medicine, and the military, but addressing environmental issues is one of the most successful GIS applications (Goodchild 1993). GIS is an ideal tool to help environmental researchers solve complex conservation issues. GIS models have been and will continue to be important tools in addressing the challenges that natural resource managers face, because models help facilitate a better understanding of the ecosystem as well as forecast impacts of environmental change (Brady and Whysong 1999).

The overlap and relationship between research and technology offer promising results (Fedra 1993). Furthermore, environmental modeling allows the integration of different fields of study to explore new avenues to address an issue. A well-tested model is a strong representation of the environment as a whole, its dynamics, and its responses to possible changes (Peng *et al.* 2002). The basic concepts of GIS deal with location, spatial distribution, and relationships, with spatial objects being the primary elements. In environmental modeling, spatial objects are specific entities that may be biological, chemical, or environmental media such as air, water, or sediment. The basic concept is expressed as the state or condition of the entity, expressed in terms of numbers, mass, or energy.

Green (1999) stated that resource management has always required the use of models to describe the resources of interest, to predict what will happen to these resources if certain actions are taken, and to prescribe the best course of action according to the specified goals. Spatial pattern analysis is in continued demand as all ecological processes happen in a spatial context (Klopatek and Francis 1999). The spatial component makes GIS the optimal platform for the analysis of process and distribution. Identifying and quantifying spatial patterns in the distribution helps researchers understand the biological importance of a species within its community.

Environmental issues are dynamic in nature. A successful GIS can be adapted to various conditions and maximizes its potential through the integration of existing datasets and dynamic variables (O'Conner *et al.* 1999). The system must be able to handle updated information, naturally occurring phenomena, and higher resolution data as they are developed. Environmental models face issues of uncertainty due to their dependency on data of variable quality gleaned from a wide variety of sources. Their need for explicitly spatial, as well as temporal, dimensions increases their complexity (Couclelis 2002).

Both modeling and GIS are important contributors to natural resource management because GIS provides a powerful tool for storage and analysis of spatial data, and the resulting models facilitate organization of knowledge about ecosystems (Brady and Whysong 1999). Falconer and Foresman (2002) outlined five basic steps in visualizing geographic understanding:

1. Formulate and ask spatially relevant questions.

2. Locate and acquire spatially relevant data and information sources.
3. Harness the geographic elements of data and explore their meaning.
4. Conduct geographic information analysis.
5. Generate visual products that act upon spatial reasoning and understanding.

These steps are the basis for the construction of any geographic model, including environmental models. They offer a clear guide to understanding the problem or issue at hand and the steps for generating beneficial results. GIS models provide the tools for beginning to comprehend temporally complex and spatially dynamic elements of our environment.

2.2 Wildlife Biology and GIS

Spatial wildlife biological data are a crucial element in managing wildlife conservation efforts around the world. Of concern to conservation biologists and wildlife managers are questions that involve predicting the future of endangered and threatened species (Akcakaya 1994). Some examples of such wildlife conservation questions are as follows:

- What is the spotted owl's chance of recovery from its current threatened status?
- What is the Florida panther's risk of extinction in the next 50 years?
- Is it better to prohibit hunting or to provide more habitats for African elephants?

- Is captive breeding and reintroduction to natural habitat patches a viable strategy for conserving black-footed ferrets?
- Is it better to preserve a large fragment of old-growth forest, or several smaller fragments of the same total area?
- Is it better to add another habitat patch to the nature reserve system or enhance habitat corridors to increase dispersal among existing patches?

At the heart of each of these questions and most questions in wildlife conservation are species occurrence data. Where are the species in question located? Why are they located there? Population ecology is concerned with how populations of plants, animals, and other organisms interact with their environment and is based on an understanding of how these populations change from one place to another (Akcakaya *et al.* 1999).

A GIS visualizes spatial data and identifies patterns in wildlife studies. In recent years, biological studies have sought to describe relationships between animal locations and habitat types. Wildlife ranges can be quickly determined in a cost effective manner through the use of a GIS (Brehme 2001). Vast amounts of data can be quickly and efficiently processed and used to facilitate decisions concerning the fate of a species. The GIS provides a visual platform for more complex analyses.

2.2.1 Case Study: Adverse Effects of Acid Rain on the Distribution of the Wood Thrush *Hylocichla mustelina* in North America (Hames *et al.* 2002).

The wood thrush (*Hylocichla mustelina* Gmelin) breeds in habitats characterized by a wide range of acid deposition values, which makes it an optimal target species for investigating how acid rain affects forest birds. This study used multiple logistic regression to test for the adverse effects of acid rain on breeding wood thrush populations, while also controlling for regional abundance, landscape-level habitat fragmentation, elevation, soil pH, and vegetation. The combination of these data layers within a GIS presents strong evidence that implicates acid rain in the decline of northeastern forest birds.

Data were imported into ESRI's ArcView GIS platform and reprojected into a common datum and coordinate system as necessary. Point data collected for factors such as soil pH and acid ion deposition were interpolated to provide estimates for each variable at unsampled locations. The data were then spatially joined with the wood thrush point data, and the combined dataset was used to fit a multiple logistic regression model. The response variable was the presence/absence data gathered for the wood thrush and the predictor variable was the estimated amount of acid deposition at each site. Covariates included landscape-level forest fragmentation, vegetation, soil pH, and the abundance of wood thrush present at each site. The model used a manual backwards elimination process to reject uninformative predictors. Predictors were retained if the probability value for the Wilks χ^2 was > 0.10.

The regression produced significant results and showed an adequate relationship between elevation, vegetation, and forest fragmentation. The model indicated a strong negative effect of acid deposition on the predicted probability of the wood thrush. The model predicted a reduced probability of breeding at high elevations and under low tree canopy and an increase in probability with an increase in forest patch size. The probability of breeding is further reduced with the interactions of acid deposition and fragmentation, acid deposition and sapling density, and acid deposition and elevation. The results suggest that atmospheric wet deposition of acid ions plays a major role in the recent decrease of some birds, including the wood thrush in the eastern United States, especially in high elevation areas with low soil pH.

2.2.2 Case Study: Using a GIS Model to Assess Terrestrial Salamander Response to Alternative Forest Management Plans (Gustafson *et al.* 2001).

The objectives in this study were to construct a GIS model to (1) predict salamander abundance in Hoosier National Forest (HNF), Indiana, (2) test the model by sampling to determine salamander abundance across the study area, and (3) use the model to compare the predicted response of salamanders to different timber management practices within HNF. Salamanders were chosen as the target species because they are a sensitive species and responsive to changes in the forest environment. This study provided an example of how GIS analytical tools provide resource mangers an avenue to evaluate potential ecological impacts of management alternatives.

The predictive salamander habitat model was developed in the GIS using known characteristics of salamander habitat, such as soil moisture, slope position, and stand age. The GIS output was verified by collecting empirical data on terrestrial salamander abundance. Salamander abundance at each of the 20 sites surveyed was plotted against the model prediction and the comparison was fit with a regression line for verification. The prediction model was ready to be tested against the land management strategies.

Five land management alternatives were simulated for each site based on previously designed management plans. The predictive salamander model was applied to each scenario and predictive salamander abundance maps based on the management practices were created. Relative effects of harvest intensity, harvest area, and harvest rate were evaluated using a repeated measures ANOVA test. The average salamander abundance was affected more by harvest intensity rather than harvest area. Harvest intensity accounted for 68–76% of the variability of salamander abundance, while harvest area and harvest rate each accounted for 4–16% of the variability. Abundance was correlated to forest age, and lower harvest rates resulted in an increased mean age of the forest. The management scenario that used the harvest intensity from a forest plan amended in 1991 produced the highest abundance of salamanders among the management alternatives.

2.2.3 Case Study: Least-Cost-Path Corridor Analysis Analyzing Wildlife Movement Corridors in Montana Using GIS (Craighead and Walker 1997)

The purpose of this study was to model potential regional-scale wildlife corridors between protected areas in Montana. This approach offers a biologically defensible assessment of probable, or optimal (least-cost), corridors by using GIS software combined with habitat data and habitat preferences of cougar (*Felis concolor* L.), elk, and grizzly bear (*Ursus arctos* L.). Suitable habitat and human interactions within that landscape were used as the two general factors that affected wildlife movements.

The study made the following four assumptions to best model the least-cost corridor route:

- good corridors are composed primarily of preferred habitat types
- humans pose problems for successful transit

- current human developments are permanent
- the least-cost path, based on the previous assumptions, offers an animal the greatest probability of survival while in transit between protected areas.

Vegetation type and the length of forest along the shrubland border were input into the GIS for a measure of habitat quality, and road density represented human activity in the area. These three layers were reclassified based on the suitable living conditions for the different species and then synthesized into a single coverage. The resulting grid represented the degree of resistance, or cost, associated with movement across a particular cell.

A cost surface was generated for each species and added to a GIS layer that contained protected areas in Montana to generate least-cost travel grids. The best corridors indicated between the protected areas varied by species due to different habitat preferences that were present or absent within an area. The results contained grids of probable transit routes for wildlife in Montana, and identified barriers and bottlenecks to the wildlife movements. The results showed that there are available corridors between land conservation areas, and they help land managers conserve Montana wildlife by keeping these areas from development. As the land becomes more fragmented, these models will assure that wildlife will have available corridors to move between available habitat areas.

2.3 Habitat Suitability Models

Habitat suitability models are an integral part of biological conservation efforts. Sampling can be more efficient with the use of a habitat suitability map, and relationships between a species and the environment in which it lives are more easily evaluated. The best way to manage specific wildlife populations is to understand why they are found where they are found. By understanding the habitat requirements of a species, a better knowledge of its distribution can be attained. Knowing where something is found and understanding how its characteristics and interactions with its environment are determined by its location is the foundation of geographic thinking (Falconer and Foresman 2002). The identification of habitat variables that are both readily available over large areas and that are correlated to species occurrence is essential to effective management of large areas (Shriner 2001). A combination of several environmental factors may play a role in the suitability of a particular habitat and therefore the habitat can be defined in multidimensional space. These dimensions make up the fundamental ecological niche of a species, which is a critical determinant of its overall distribution (Peterson 2001).

Modeling the habitat for specific organisms targeted for conservation give planners and decision makes a tool to assist them in optimizing biodiversity preservation (Egbert *et al.* 1999). Predicted distributions for a species based on known habitat characteristics can be effectively used for this purpose. Maps are typically developed by filtering and combining key features of the environment that are correlated to the known habitat requirements of the species (Shriner 2001, O'Conner *et al.* 1999, Egbert *et al.*

1999). Once these habitat indexes are developed, they can be used to track the changes in habitat availability (Egbert *et al.* 1999, Peterson 2001). Mapping a habitat within a GIS can reveal the nature of conflict between wildlife and human populations. Decision support for managers is among the essential functions of habitat models and other geographic analyses (Falconer and Foresman 2002).

Even if the influence of environmental factors on a species' distribution is not known, a GIS can be used to evaluate the spatial coincidence of populations to their surroundings. Johnston (1993) noted that this type of analysis not only provides relationships that can be used in models that infer population numbers based on habitat characteristics, but it also provides information on the spatial extent and overlap of home ranges that can be used to analyze interactions among individuals in a population. The GIS provides an informed study platform that can be verified by ground truthing. If the GIS is correct, a visit to any positive location will confirm the information mapped for that particular spatial extent (Falconer and Foresman 2002).

2.3.1 Case Study: A GIS Analysis and Model of Suitable Gray Wolf Habitat in the Northern Rockies (Brehme 2001)

A GIS was used in this study to better understand the habitat requirements of the gray wolf (*Canis lupus* L.) in central Idaho. The goal was to restore sustainable populations in ideal habitats throughout its historic range. Most researchers agree that the two most important characteristics of grey wolf habitats are an absence of humans and adequate prey. The use of GIS for the habitat model provided a fast, cost-effective method for locating and identifying the potential range of the grey wolf. Large amounts of data are efficiently managed, which helps facilitate timely decisions about the fate of the species. The specific objectives of the study were to:

1. Examine the currently occupied wolf home range in Idaho and western Montana and determine what landscape-scale characteristics were associated with wolf presence.
2. Incorporate the most significant habitat characteristics in a GIS model that would predict potential wolf habitat throughout the study area.

Wolf data were collected by using global positioning systems (GPS) telemetry collars over a 22 month period. Ten pack areas were compared to ten non-pack areas to test for significant habitat characteristics. Data sets were included that represented habitat variables, such as slope, aspect, elevation, ownership, and land cover. A United States Geologic Survey (USGS) Digital Elevation Model (DEM) was used as the elevation variable and to generate the slope and aspect grids across the study region. The mean slope, aspect, and elevation were calculated for each of the pack and non-pack areas as well as the percentage of land cover and ownership type for each.

The data were then entered into Statistical Analysis Software (SAS) in a stepwise logistic regression analysis. The logistic procedure correctly classified 70% of the pack and non-pack areas ($P < 0.05$). To create the habitat maps, the logistic equation was entered into Arc/INFO GRID. The results were applied to each significant grid layer and the output resulted in a grid of cells ranging between 1 and 0, representing the probability of presence or absence. The analysis proved to be effective and adequate for small sample sizes. It was found that 89% of the predicted habitat areas were composed of

Forest Service lands and just over 6% were composed of private lands. Much of the suitable habitat area has yet to be visited by wolves and the outlook for wolf recovery is positive.

2.3.2 Case Study: A GIS-based Habitat Model for Wood Thrush, *Hylocichla mustelina*, in Great Smoky Mountains National Park (Shriner *et al.* 2002)

Knowledge of habitat preferences and ecology is required for proper management of wildlife populations. The use of wood thrush as a study species has been the focus of several studies that investigated the impact of fragmentation and land-use change on population abundance, distribution, and productivity. The purpose of this study was to investigate the use of commonly available GIS data to assess different topographic indices and habitat characteristics to predict the presence or absence of wood thrush species within GSMNP.

Over a span of four years, more than 4,000 circular plot point counts were conducted for the presence or absence of the wood thrush species. Points were established systematically along the established trail system and roadways, located 250 meters apart to avoid double counting, and were collected using a Trimble GeoExplorer II GPS unit. They were stratified throughout the park with respect to the different vegetation types. Habitat characteristic data included vegetation type, bedrock geology, and disturbance history.

The topographic indices included elevation, slope, aspect, topographic convergence index (TCI), terrain shape index (TSI), landform index (LFI), Shannon-Weaver Index of Topographic Complexity (SWI), and topographic relative moisture index (TRMI). TCI is the potential soil moisture developed to simulate runoff saturation and infiltration. The TSI was developed to distinguish between ridges/exposed areas and coves/protected areas. The LFI describes general protection classes at a site such as cove, slope, or ridge. The SWI is an index of topographic diversity. The TRMI was developed to describe local moisture regimes.

The data were processed through a logistic regression model that returned the probability of detecting a wood thrush based on the habitat and topographic variables. The topographic variables were standardized prior to analysis and the 13 vegetation types, 23 bedrock geology types, and five disturbance classes were treated as class variables. Only type variables that were represented by a minimum of 30 observations were included. The final analysis was performed on 10 vegetation types and 18 geology types. The researchers used backward elimination and forward selection processes, which analyzed the full model and then removed variables one at a time if they failed to meet the specified significance level for the model. The SWI and LFI were found to be topographic indexes significantly associated with the presence of the wood thrush as well as elevation, disturbance history, and geology type. All variables included in the model were significant ($P < 0.05$). Eighty-three percent of the observed data points were correctly identified. Each explanatory variable was queried within the GIS to determine the value for each cell. Each cell received a probability of detection value calculated by the logistic regression equation. The result was a 90 x 90 meter grid map that assessed the probability of detection of wood thrush throughout GSMNP.

2.3.3 Case Study: A GIS Model for Identifying Potential Black-Tailed Prairie Dog Habitat in the Northern Great Plains Shortgrass Prairie (Proctor *et al.* 1998)

The prairie ecosystem including the black-tailed prairie dogs (*Cynomys ludovicianus* Ord.) depends on the identification and conservation of the remaining suitable habitat areas for the species. The purpose of this study was to provide a methodology for creating habitat maps based on prairie dog interactions with specific environmental characteristics so that conservation efforts on the Charles M. Russell and UL Bend National Wildlife Refuges and similar lands could be prioritized.

Vegetation, slope, and soils data of the area were combined with known prairie dog communities mapped by a GPS unit to examine whether black-tailed prairie dogs choose:

- short to medium grassland cover types more than expected
- slopes of 0–8% more than expected
- soils ranging in texture from clay to sandy loam more than expected
- soils with depths greater than 60 inches more than expected

The expected levels for each category were chosen after a thorough review of the literature on related studies, expert interviews, and spot checks of several known colony locations.

All layers used in the study were combined into one dataset. The North dataset contained 517 unique combinations of the variables while the South dataset contained 316. A classification tree was used to find which variables were the most strongly associated with the presence of the species. A logistic regression model was also used to further analyze the datasets to explain the variation between available and occupied habitats. The North dataset classification tree showed that prairie dogs used four specific vegetation types. Slopes greater than 4% were not used by prairie dogs, and clay-loam soils were the only significant soil type used by prairie dogs. The South dataset contained less information. The logistic regression model showed similar results to the classification tree with slope being the single most important factor for prairie dog presence. The North dataset classification tree was used to derive the habitat suitability model. About 10% of the area was classified as suitable habitat with only 1.2% of that being classified as optimal habitat area.

Prior studies found that soils had a high correlation with the presence of prairie dogs, but this newer model suggested soils had little value to the distribution of prairie dogs. Previous studies also investigated possible relationships with soil classes, slopes, and certain vegetation types. Vegetation types associated with prairie dog presence all share the common characteristic of being low-biomass coverages, which may be the cause of the strong relationship.

The habitat suitability map can be used to help managers predict where future expansion of the black-footed prairie dog will occur either on their own or through reintroduction efforts. A large area of the preferred habitat remains unoccupied by the species. This area falls on private land and may be subject to poisoning by the landowners. Patterns among core areas of prairie dog occupation and possible connecting corridors can be observed and can aid managers as they prioritize the conservation of certain areas to help benefit the prairie dog ecosystem on a large scale. Other management implications for this model include improved black-footed ferret reintroduction, development of a plague management plan, development of a prairie dog ecosystem management plan, and development of a prairie dog ecosystem conservation strategy.

2.3.4 Case Study: Prediction of Thistle-infested Areas in Badlands National Park Using a GIS Model (Price and Tinant 2000)

This study analyzed the spatial correlation between Canada thistle (*Cirsium arvense* L.) and environmental parameters within the Burns Basin area of Badlands National Park. This invasive, noxious weed is found throughout the park. It potentially threatens the black-footed ferret by harboring its predators, displaces native plants, and economically degrades adjacent agricultural lands. Field identification of problem areas is costly and time-consuming. A predictive GIS model could allow managers to make effective decisions on how to use herbicides to defend against the invasive plant.

A logistic linear regression model was applied to slope, distance to streams, distance to water bodies, distance to wetlands, distance to roads, Normalized Difference Vegetation Index (NDVI), and Normalized Difference Moisture Index (NDMI). NDVI is a measure of "greenness" and in this study it roughly correlated with biomass. NDMI is a representation of soil moisture. Horseback surveys were conducted at 280 different points to gather spatial information on thistle-infested areas and thistle-free areas. Logistic regression established relationships between the thistle data and the environmental factors. Slope, NDVI, and NDMI were shown to be the best predictors of thistle growth from the regression analysis. Next, environmental variables, such as geology, soil type, aspect, and vegetation types, were analyzed in a prediction model.

Aspect was not found to be significant, but highly significant positive and negative correlations were found for geology, soil, and vegetation classes.

Each method produced a separate prediction map, but to integrate both models the two maps were combined. The regression map was weighted to equalize the values of the two maps. The result was a probability map that showed prediction based on all tested significant environmental factors. Field tests on separate transects in the park were next conducted to assess the accuracy of the model. Points along each transect were classified as thistle or non-thistle. Locations where thistles were found all fell within high probability areas. However, some points previously surveyed fell within low and moderate probability of occurrence areas. The models provided habitat suitability predictions rather than probability of actual thistle occurrence. The combined map provided the highest accuracy of the three models. The final model is helping managers estimate the total acreage of the park that is susceptible to thistle infestation.

2.4 GPS - GIS Integration for Environmental Research

Global positioning system (GPS) receivers collect spatial data that can be modeled within the framework of a GIS. GPS technology provides information about absolute positioning, relative positioning, and timing for a wide range of applications. Environmental modeling and decision support systems have been improved by the integration of GIS and GPS technologies (Varma 2002). GPS units provide a quick means for capturing coordinates and attributes of features in the field (Falconer and Foresman 2002). Having this seamless process in place is valuable for studies that require a large amount of data be entered into the GIS. GPS technology makes it easy to construct and maintain an accurate GIS database. The technology has become a valuable tool in environmental research, a trend that is sure to continue in the coming years. The ease of obtaining different kinds of information at once and later evaluating the data in a GIS make the technology a vital part of wildlife population studies (Blake and Brown 2001).

2.4.1 Case Study: Mapping Southern Oregon's Noxious Weeds with GPS (Kiser and Williams 2000)

This study incorporated GPS and GIS technology for fighting plant invasions. The emergence of these geospatial technologies provides land managers an opportunity to develop a standardized noxious weed database that could carry into the future and help create efficient control strategies. Previous mapping of the area's noxious weed occurrences consisted of highlighted or circled areas on a district map, and different managers had different methods of recording observations. GPS and GIS mapping systems not only allow for better mapping efforts, but also provide better structure of data gathering and reporting.

GPS units capable of 75 cm accuracy were used to map all previously located and newly discovered noxious weed populations. Invasive plant communities were identified on aerial photographs to save time in the field and small-scale digital photographs were taken on site. GPS points and attributes about invasive plant populations were collected in the field and were imported seamlessly to the GIS for further analysis. The digital photographs were also linked to the points within the GIS for a quick visual reference of the area in question. The system allowed managers to map new populations and monitor changes in existing sites. Strong correlations were found between presence of a weed community and road shoulders. Roadways with heavy use have greater weed populations, and it appeared that vehicles and heavy equipment are the major cause of weed seed dispersal.

2.4.2 Case Study: Mapping Urban Deer Populations using GPS and GIS (Blake and Brown 2001)

The purpose of this wildlife survey was to document various factors that influence white-tailed deer (*Odocoileus virginianus* Zimmermann) communities in Hollywood Park, Texas. Data were collected using a GPS receiver and a laser rangefinder. A data dictionary was used to predefine attribute menus for easy data entry in the field. Defined attribute data included antler presence/absence and distance to a feeder house, which were entered at each deer occurrence GPS point.

The study area was driven morning, afternoon, and evening to determine temporal patterns of the deer. When a deer was spotted, a point was taken and the laser rangefinder was used to "tag" the deer to calculate its exact position. Deer data were overlaid with a roads layer, and feeder house data in the GIS. Two hundred meter buffers were created around feeder houses for further analysis. ANOVA tests were conducted to determine if a statistical difference existed in the number of deer counted at different times of day. The evening count was found to be statistically different from the morning and afternoon counts and could be attributed to more deer feeding in the evening as well as the nocturnal nature of deer. An ANOVA test was also conducted for deer counts within the

200 m buffer of the feeder houses. No statistical difference between times of day and locations of the deer was found. Deer were always observed within the 200 m buffer zone. The data collected and analyzed in this study could help managers determine how to best manage the overabundance of deer in urban areas.

2.4.3 Case Study: GIS Mapping Aids Understanding of Elephant Seal Behavior (Galimberti and Sanvito 1999)

Temporal GIS-GPS studies are another beneficial use of the technology. Researchers in the Falkland Islands used GPS mapping to help understand southern elephant seal (*Mirounga leanina* L.) behavior. The main objectives of this study were to use GPS and GIS technology to identify variation between individuals' mating tactics and breeding strategies and to obtain data about the general breeding biology of the population. Nine to twelve harems were spread over the 7 km^2 area and each harem's size and shape were mapped by walking the perimeter with a GPS unit. Data were also collected on individual males outside the harem, as well as isolated females. By collecting data every few days, researchers were able to model within the GIS how the dominant male changes the shape of the harem to protect against individual males outside of the harem. The ability of peripheral males to gain access to harem females depends on the area and shape of the harem and the density of females in the harem. Using a GIS, researchers observed that different harems have different shapes based on environmental properties such as topography. Densities of females depended on the capability of the harem master to keep females packed together.

Chapter Three
Study Area Description

3.1 Great Smoky Mountains National Park

Great Smoky Mountains National Park (GSMNP) was established by Congress in 1934 and comprises over 308 km^2 of wilderness along the Tennessee-North Carolina border (Fig. 3.1). Willis P. Davis of Knoxville, Tennessee, officially started the movement for a National Park in 1923; state funds were matched by the Laura Spelman Rockefeller Memorial and land acquisition began. Through combined gifts, private donations, and public funds, the Governors of Tennessee and North Carolina presented 64,294.86 hectares of land to the U.S. Secretary of the Interior in 1930. President Franklin D. Roosevelt dedicated the park on September 2, 1940, at Newfound Gap, with the mission of preserving and protecting the wild beauty and natural features contained within the borders of the park (Pierce 2000).

The park is known for its extraordinary biodiversity. Over 1,500 kinds of flowering plants alone are found in the park, which is more than any other national park in the United States and Canada (White 1996). Given its beauty and location, the Smokies have become the most visited national park in the nation with over 9 million visitors per year in recent years.

The park is the second largest park in the eastern United States. It contains over 124 km of streams and 528 km of maintained trails, most of which follow the ridgetops and stream corridors, because these locations make for easier maintenance and cause less damage to the park's terrain. The trail system covers most of the geographic regions

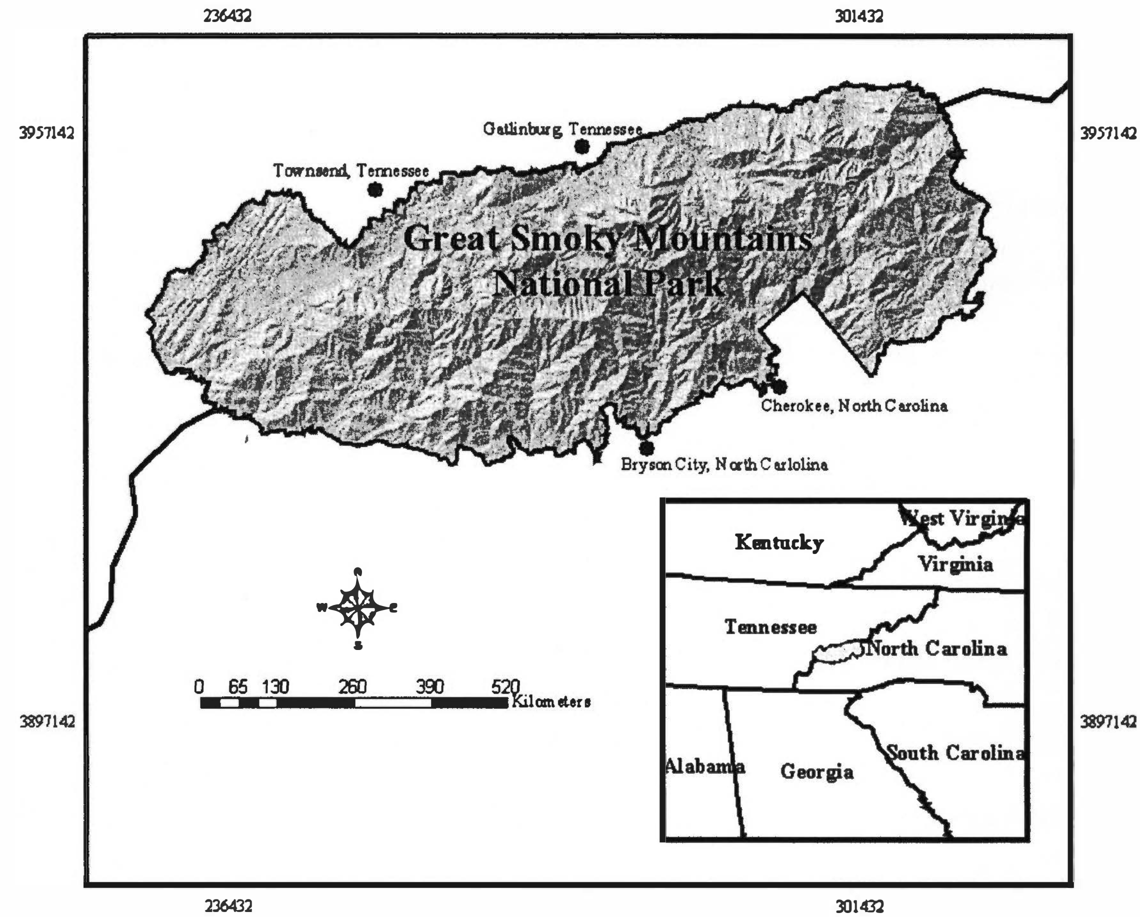


Fig. 3.1. Great Smoky Mountains National Park. The park is located astride the eastern border of Tennessee and the western border of North Carolina (ESRI and NPS data).

found in the park (GSMNP 2001). Over 70% of the land area is forested, providing the largest extent of forested landscape in the eastern United States (Shriner 2001). Elevation in the park ranges from a low of about 204 m in the park's southwest corner to Clingman's Dome, which at 1,830 m is the third highest peak in the eastern U.S. (White 1996). The altitudinal range resembles the latitudinal extent one would experience traveling south to north from Georgia to Maine. In the lowlands, trees typical of the South, such as sweetgum (*Liquidambar styraciflua* L.), are in abundance while higher elevations contain species like American mountain ash (*Sorbus americana* Greene) normally found in northern habitats. Habitats vary greatly by elevation as well as with factors such as slope and aspect. A habitat on a dry, sunny, south-facing ridge is very different from one found on a moist, cool, northern slope even if the two are located at the same elevation (Kemp 1993).

3.2 Geology of the Great Smoky Mountains

Geology has a profound influence on all life in the park due to the processes of geochemistry and geomorphology. The geologic structure determines how and at what rate surfaces change. The chemical makeup of the geologic foundation also has a direct effect on the soil and stream characteristics.

Sedimentary rocks are the youngest rocks found in the park and include sandstone, limestone, and siltstone. Igneous rocks include granite and basalt and are seldom encountered in the park (Southworth 2000). Older rock material is more likely to be changed, and because many of the rocks in GSMNP are ancient in origin, metamorphic rocks are common (Wuerthner 2003). Igneous and metamorphic rock types can be found in the southern and southeastern sections of the park, but the dominant rock type of the park is metamorphosed sedimentary rock (Southworth 2000). The most dominant rocks of this type found in GSMNP are the Thunderhead and Anakeesta Formations (Fig. 3.2).

3.3 Soils of the Great Smoky Mountains

Ultisols and inceptisols are the two main soils found in the region (Stephensen *et al.* 1992). Ultisols are defined in part by their low contents of weatherable minerals and are susceptible to acidification by acid deposition unless there is sufficient buffer capacity (Binkley 1989). The nitrates deposited in the high elevations of the park are six to seven times the amount that local soils can naturally process. The release of bases into the soil is equal or less than the removal by leaching, with most of the bases being held in the upper few centimeters of the soil as well as in the vegetation. To form, ultisols require a mean annual temperature of greater than 8°C (NRCS 1999).

Inceptisols are found on fairly steep slopes, on young geomorphic surfaces, and on resistant parent materials (University of Idaho 2002). Inceptisols are typically mineral-rich. Water is available in the soil for well over half of the year (NRCS 1999).

The soils in the park have been classified into 15 specific soil units (Table 3.1). Over 90% of soil units 1, 2, 5, 9, and 10 comprise upland soils, landslides, and heath balds. These soil units are very low in plant nutrients. The soil depth for these upland soils ranges from deep to moderately deep as well as a few shallow classes. Upland soils comprise between 70% and 89% of the soils in units 3, 4, 6, 7, and 8. Most of these soils

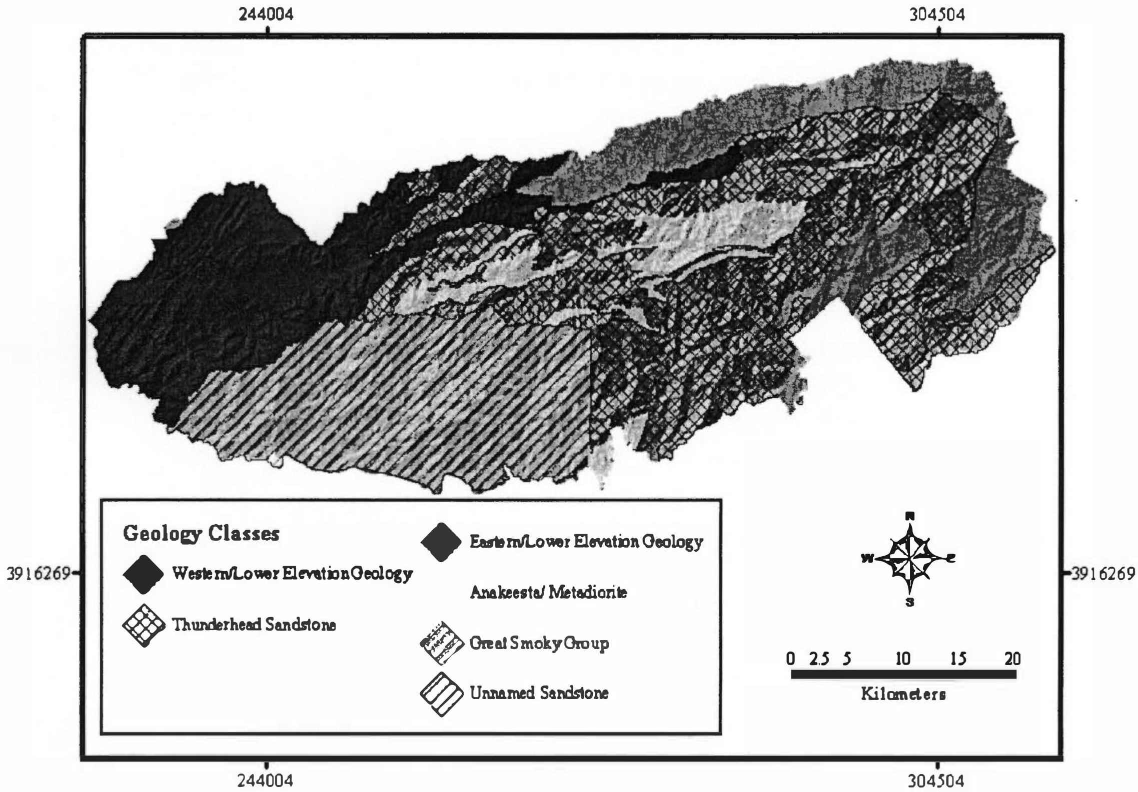


Fig. 3.2. Geology of GSMNP (data obtained from the NPS).

Table 3.1. Soil Units of GSMNP (data obtained from the USGS)

Name	Class
Luftee-Anakeesta soils	Unit 1
Breakneck-Pullback soils	Unit 2
Oconaluftee-Guyot-Chiltoskie soils	Unit 3
Wayah-Tanasee soils	Unit 4
Cataska-Sylco-Spivey soils	Unit 5
Ditney-Unicoi-Spivey soils	Unit 6
Soco-Stecoah-Spivey soils	Unit 7
Evard-Cowee-leaterwood soils	Unit 8
Cataska-Sylco-Tsali soils	Unit 9
Junaluska-Tsali soils	Unit 10
Spivey-Santeetlah soils	Unit 11
Lauda-Fanning soils	Unit 12
Dellwood-Smokemont-Reddies soils	Unit 13
Lonon-Cades-Allegheny-Rosman soils	Unit 14
Junaluska-Brasstown-Spivey soils	Unit 15

are also low in plant nutrients and the soil depth varies greatly. For most of these soils, aspect does not cause a change in the surface soil morphology in the shaded well-formed head slopes. Many of the soils in these units have both frigid and mesic temperature regimes. The typically deep soils found in the lower elevations are units 11, 12, 13, 14, and 15. Some of these soils are found in or on the fringe of the floodplain and are typically well drained (Fig 3.3).

3.4 Climate of the Great Smoky Mountains

The climate of the park encompasses conditions from warm to cool temperate, but rainfall is abundant throughout the park as maritime tropical air masses from the Gulf of Mexico bring year-round moisture to the Smokies. Typically, plant diversity increases from cooler parts of the world to warmer ones and from drier regions to wetter ones. The Smoky Mountains fall midway along the global temperature gradient and are on the wet end of the moisture gradient. The lack of a dry season helps promote diversity in the park, making it the most biologically complex in all of North America. Generally more rainfall occurs in the higher elevations, which receive an average of 254 cm, while the lower elevations receive an average of 140 cm (White 1996).

The park climate is seasonal. Stable high-pressure systems from the southern North Atlantic Ocean create dominant sunny weather in the summer months. Afternoon thunderstorms are common occurrences in the hot and humid summer months and temperatures range from 35°C during the day to 18°C at night. In early fall, daytime temperatures fall to between 21°C and 28°C. Fall is also the driest time of the year in the

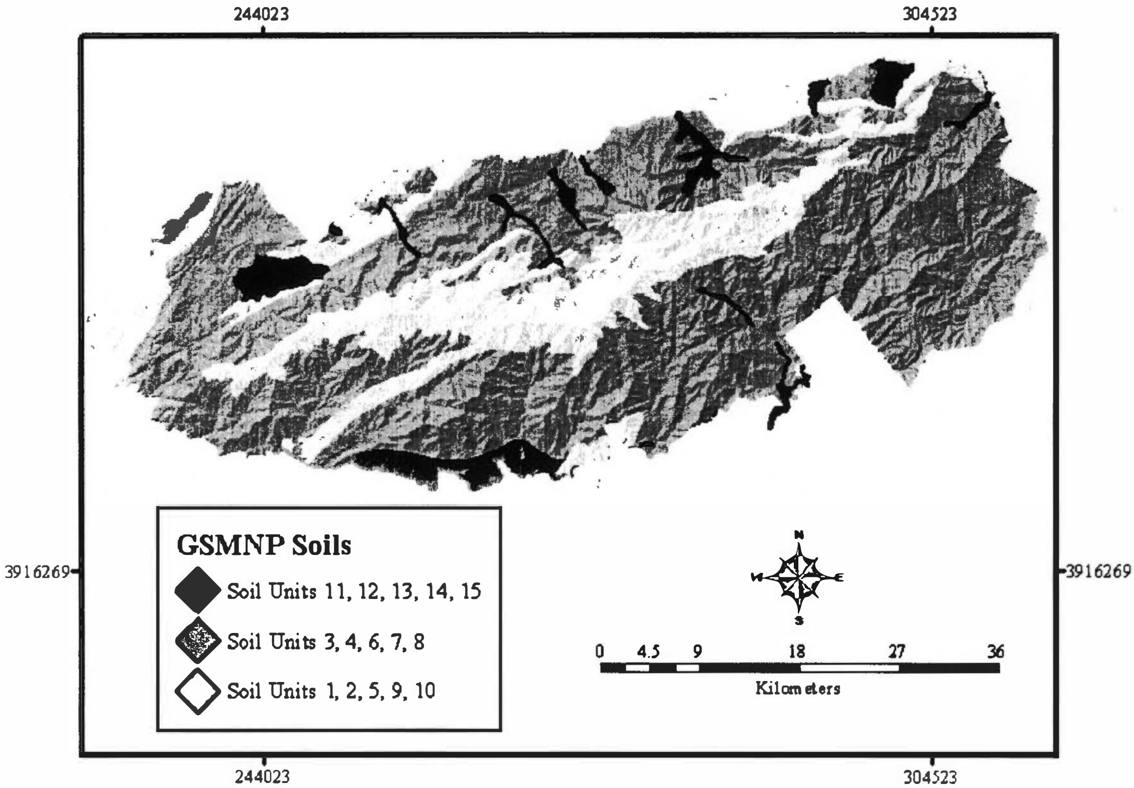


Fig. 3.3. Soil classes of GSMNP (data obtained from USGS).

park. Freezing temperatures start to occur by November.

The winter climate in the Smokies is influenced by storms rising from the Gulf of Mexico and interacting with the western slopes of the mountains. This weather is compounded by storms arriving from the southern North Atlantic that encounter the eastern slopes. In contrast to the short thunderstorms of the summer months, winter storms often last days. Daytime temperatures during the winter months range between 10°C and 15°C in the lower elevations (Table 3.2) and below freezing in higher elevations (Table 3.3). Snow is common in the high elevations of the park, especially in the spruce-fir zones. The frost-free season ranges from 200 days in lower elevation areas to 100 days in the higher areas (White 1996). By mid-April, temperatures start to rise again reaching 28°C with lows down to 5°C. Rainfall in the springtime (May–April) averages between 10.1–11.4 cm (Wuerthner 2003). Throughout the park, summer months tend to be both the hottest and wettest of the year, but the most unpredictable weather arises in the winter.

The height and structure of the mountains, along with predominant weather patterns, trap and concentrate air currents as they enter the southern Appalachians (Renfro 2001). The interactions between the rugged topography and elevation with the climate patterns of the park are the two most important factors that influence soil moisture. As elevation increases, temperatures fall and the precipitation levels increase. The average rate of temperature decrease is 1.5°C per 305 m (Stupka 1960). The added rainfall classifies certain small sections of the park as temperate rainforest (GSMNP 2001).

Table 3.2. Monthly weather at Gatlinburg, elevation 446 m (Wuerthner 2003)

Month	Average High Temperature (°C)	Average Low Temperature (°C)	Precipitation (cm)
January	10.5	-2.2	12.2
February	12.2	1.7	12.2
March	16.1	1.1	13.5
April	21.7	5.6	11.43
May	26.1	10.0	11.43
June	30.0	14.4	13.2
July	31.1	15.0	14.5
August	30.6	15.6	13.5
September	28.3	12.8	7.6
October	22.8	6.1	7.9
November	16.1	1.0	8.6
December	11.1	-2.2	11.43

Table 3.3. Monthly weather at Clingman's Dome, elevation 2,024 m (Wuerthner 2003)

Month	Average High Temperature (°C)	Average Low Temperature (°C)	Precipitation (cm)
January	1.7	-7.2	17.8
February	1.7	-7.8	20.8
March	3.9	-4.4	20.8
April	9.4	6.1	16.5
May	13.9	6.1	16.5
June	17.2	9.4	17.5
July	18.3	11.7	21.1
August	17.8	11.1	17.3
September	15.6	8.3	13.0
October	11.7	3.3	13.7
November	5.6	-2.2	16.3
December	2.8	-6.1	18.5

3.5 Forest Cover within the Great Smoky Mountains

Over 130 species of trees and 4,000 other plant species comprise one of the most complex vegetation communities in temperate America (Shriner 2001). Disturbance plays a major role in the forest development. The park has the largest remaining uncut forests in the eastern United States with over 40,000 ha having been either selectively cut or left undisturbed. The old-growth forests enhance diversity and abundance between species (GSMNP 2001).

Five dominant forest types are found in the park and represent all the major forest types in eastern North America (Fig. 3.4). These five types are spruce-fir forests in the higher elevations, northern hardwood forests in the middle to upper elevations from 1,070–1,520 m, pine-oak forests on the drier ridges, hemlock forests along stream banks, and cove hardwood forests lining the valleys. The spruce-fir forest is found along the high ridgelines running the length of the park. This forest type is the most restricted as it is only found in elevations above 1,370 m and is able to withstand the cold conditions present in these areas. Spruce-fir forests occupy only about 2% of the park (Wuerthner 2003). Trees associated with the red spruce (*Picea rubens* Sarg.) and Fraser fir (*Abies fraseri* (Pursh) Poir.) include yellow birch (*Betula alleghaniensis* Britt.), pin cherry (*Prunus pensylvanicus* L.), American mountain-ash, and mountain maple (*Acer spicatum* Lam.). Due to the dense tree cover, shrubs and other plants are often absent in these zones. Catawba (*Rhododendron catawbiense* L.) and Carolina (*Rhododendron carolinianum* L.) rhododendrons, scarlet elder (*Sambucus pubens* Michx.), thornless blackberry (*Rubus* spp. L.), hobblebush (*Viburnum lantanoides* Michx.), witherod

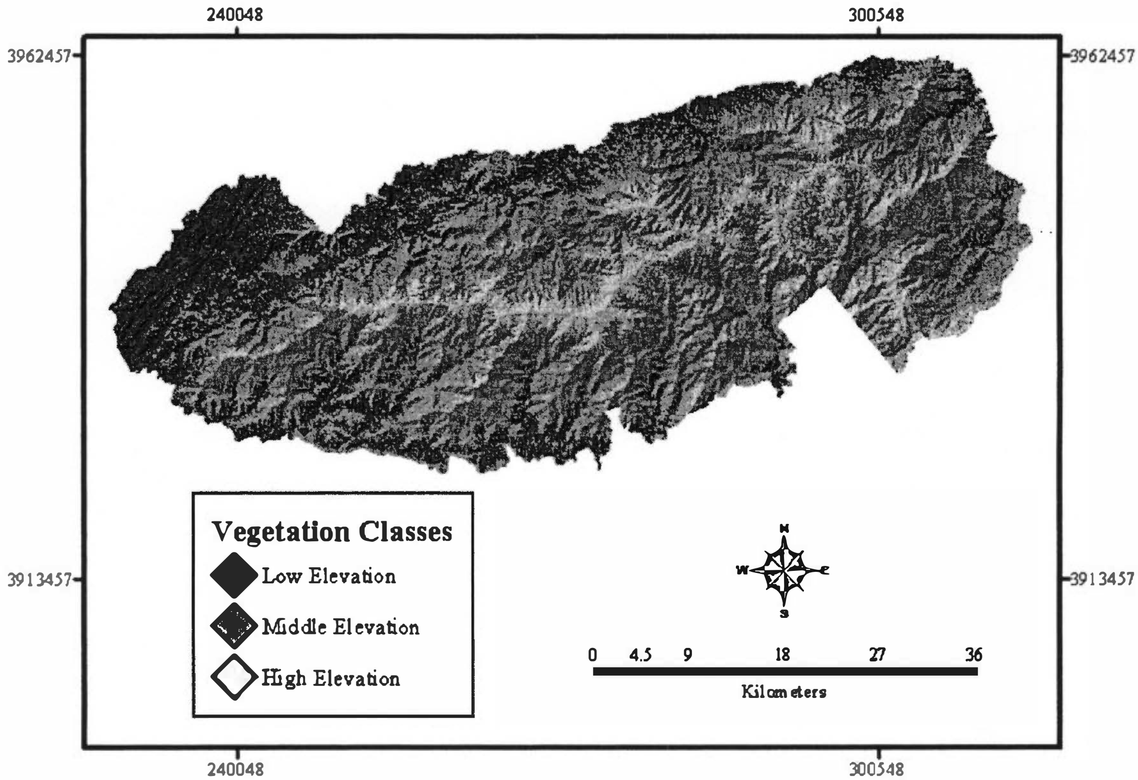


Fig. 3.4. Vegetation of GSMNP (data obtained from the NPS). Low elevation vegetation consists of pine, pine-oak, and xeric oak forests. Middle elevation vegetation consists of tulip poplar, mesic oak, mixed mesic oak, and cove hardwood forests. High elevation vegetation consists of spruce-fir and northern hardwood forests as well as heath and grassy balds.

(*Viburnum nudum* L.), and Blueridge blueberry (*Vaccinium* spp. L.) are the most common shrubs in this vegetation type (Stupka 1960).

The northern hardwood forest is largely dominated by yellow birch and American beech (*Fagus grandifolia* Ehrh.). They are often found in association with red spruce. On the lower limits of the area, eastern hemlock, sugar maple (*Acer saccharum* Marsh) and black cherry (*Prunus serotina* Ehrh.) are often present, while yellow birch, pin cherry, and mountain maple are found at the upper limits. Different tree species dominate depending on the conditions at the site. American beech is common in drier sites while yellow birch is common in wet areas. Shrubs in this forest type include smooth hydrangea (*Hydrangea arborescens* L.), catawba and rosebay (*Rhododendron maximum* L.) rhododendrons, thornless blackberry, and hobblebush. The most abundant herbaceous plants include fawnlily (*Erythronium albidum* Nutt), creeping bluet (*Houstonia sorphllifolia* Michx.), trilliums (*Trillium* spp. L.), white violets (*Viola renifolia* Gray), and yellow beadlily (*Clintonia uniflora* Schult.) (Stupka 1960).

Closed and open stand pine-oak forests are both present in GSMNP. Closed pine-oak forests are found on low to mid elevation slopes in dry areas and are dominated primarily by white oak (*Quercus alba* L.), northern red oak (*Quercus rubra* L.), and chestnut oak (*Quercus prinus* L.). Red maple (*Acer rubrum* L.), yellow poplar (*Liriodendron tulipifera* L.), and black gum (*Nyssa sylvatica* Marsh.) are also relatively abundant along with smaller trees and shrubs such as dogwood (*Cornus florida* L.), mountain laurel (*Kalmia latifolia* L.), and flame azalea (*Rhododendron calendulaceum* L.) (Stupka 1960). Open pine-oak stands are often found along ridges and rocky outcrops. In addition to the oak species found in the closed pine-oak forests, Table Mountain pine (*Pinus pungens* Lamb.), pitch pine (*Pinus rigida* Mill.), Virginia pine (*Pinus virginiana* Mill.), and chestnut oak are also associated with the open pine-oak forest. Pines dominate sunny areas such as outcrops and sandy areas, but are often overshadowed by the oaks in other areas. Wildfires are common in this forest type and are essential for maintaining several associated species (Wuerthner 2003).

Hemlock forests can be found up to 1,220 m in elevation on moist, shady slopes and can occasionally be found up to 1,520 m. This forest type is most often found along streams. Several stands of old-growth hemlocks remain in the northwest section of the park. Red maple, sugar maple, yellow birch, yellow poplar, fraser magnolia (*Magnolia fraseri* L.), and striped maple (*Acer pensylvanicum* L.) are associated with the hemlock forests. Understory shrubs include witch hobble (*Viburnum alnifolium* Marsh.) and *Rhododendron* spp. (Wuerthner 2003).

Cove hardwood forests are found in the deep rich soils of the sheltered basins and hillsides of the park. Most of these forests are found below 1,220 m. The dominant overstory trees include white basswood (*Tilia americana* L.), silverbell (*Halesia carolina* L.), yellow poplar, sugar maple, red maple, white ash (*Fraxinus americana* L.), and yellow birch, among several others. Wildflowers flourish in this type of forest from spring to early fall.

3.6 Topographic Characteristics of the Great Smoky Mountains

The Smoky Mountains present a rugged and varied topographic landscape (Fig. 3.5). Slope, aspect, and slope shape are other important topographic characteristics that help form an organism's habitat. Slope refers to the steepness or grade of the mountain, often reported as a percentage. Lower slope percentages are flatter areas, while higher percentages represent much steeper slopes. Aspect refers to the compass direction a slope faces. A south-facing slope would be represented with an aspect of 180 degrees. Slope shape refers to the convexity or concavity of a slope. These characteristics combine to influence the amount of sunlight that reaches a site, the temperature of the site, and its ability to retain moisture in the soil (White 1996). Soil moisture, a critical factor at middle and lower elevations, is heavily influenced by concavity. Rainfall along ridges and other convex landforms flows into the concave landscapes of the valleys. Therefore, convex landforms will be much drier than concave landforms. Sunlight is an important factor for soil moisture and warmth at a site. South-facing and west-facing slopes, as well as ridgelines, receive more sunlight, have higher average temperatures, and experience greater rates of evaporation. Narrow valleys in the mountains as well as north-facing slopes are in the shade for much of the day and will be cooler and moister. In the higher elevations, these topographic characteristics are less of a factor in organism distribution because there is much more moisture, cooler temperatures, and more cloud cover in these areas.

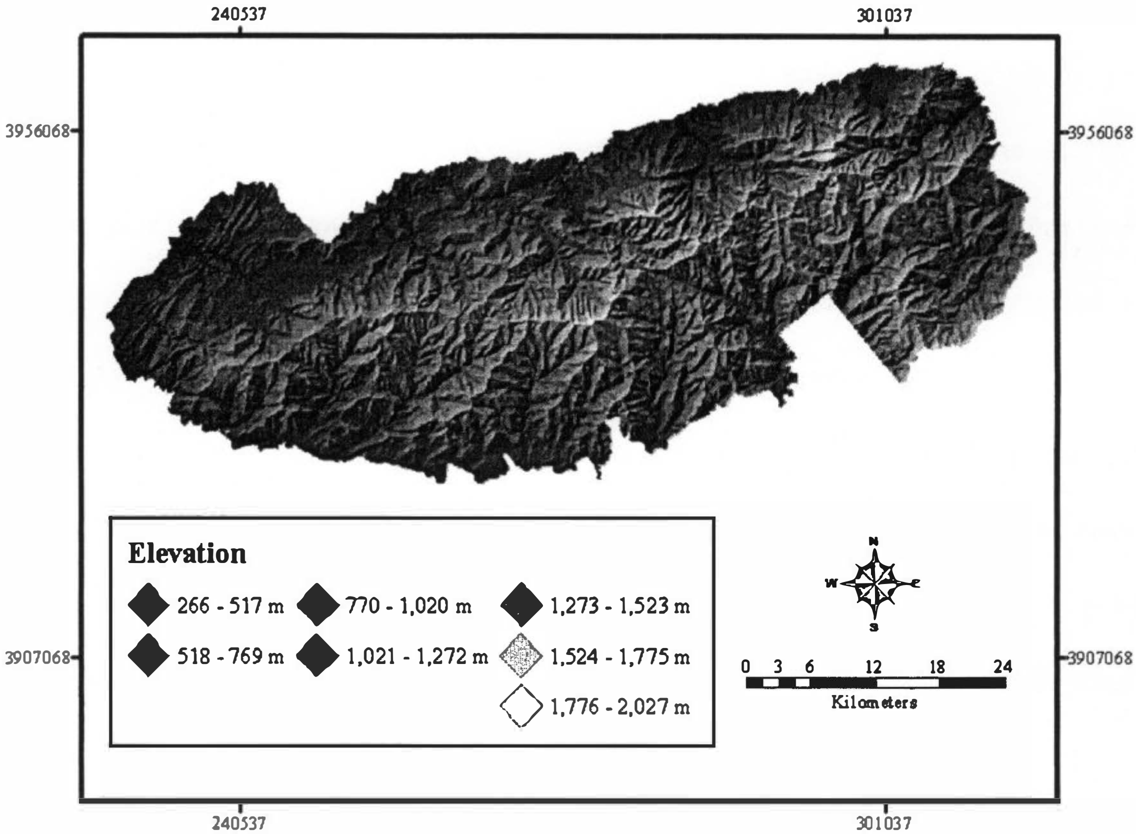


Fig. 3.5. Elevation ranges of GSMNP(data obtained from the NPS).

3.7 Disturbance History of the Great Smoky Mountains

By the early 1900s the logging industry had moved into the Great Smoky Mountains. The South was the leading timber producer from 1900 to 1920 in the United States with much of the lumber coming from the Southern Appalachian region. Before 1900 it was estimated that only 10% of the Smokies had been cleared and the area had the largest continuous forest area in all of Appalachia. Loggers cut the plentiful large sawlogs such as spruce, chestnut, and yellow poplar and floated them down the Little River to local mills. Soon railroads were constructed to aid in the timber liquidation, furthering the environmental degradation brought about by timber harvesting. By 1920 two-thirds of the park's lumber had been cut or lost to fires caused by the logging operations. Before the park was created, timber companies owned 85% of the proposed land. By the time the GSMNP was established, almost three-fourths of the land had been cut.

Many areas in the park were either clear-cut, selectively cut, or settled, but some of the lands, especially steep slopes and hard to reach areas, were left untouched (Fig. 3.6). According to the Park Service, about 36% of the park contains uncut forestlands. A remarkable difference exists between areas that had been cut and these old-growth forests. Around half of the plant species found in the old-growth forests are still missing from the disturbed areas over eighty years after the logging companies have left (Wuerthner 2003). These old growth forests have more complex structures than second growth forests of the same type, which favors higher diversity and abundances of certain species groups (GSMNP 2001). Old-growth forests offer a good picture of what the

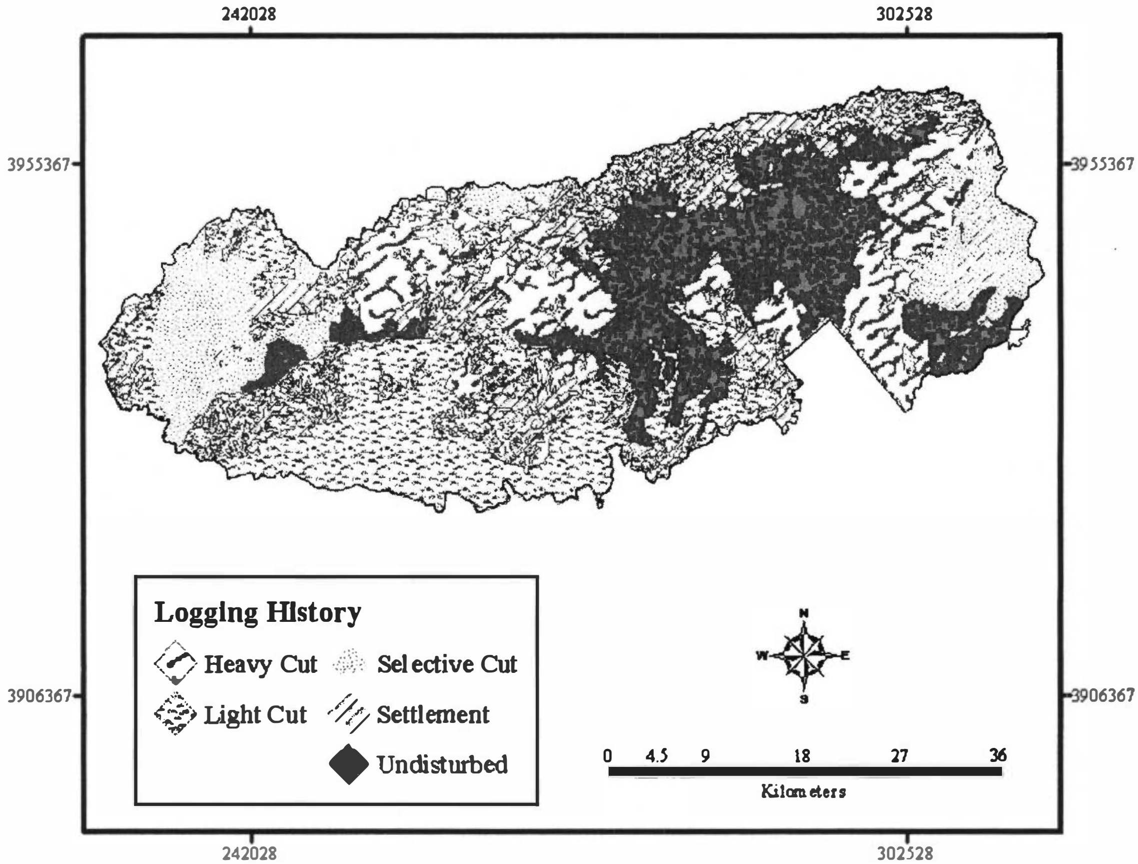


Fig. 3.6. Logging history of GSMNP (data obtained from the NPS).

forest structure of GSMNP looked like before logging and settlement came to the area.

Air pollution is another major human-influenced disturbance within GSMNP. Research conducted over the past 20 years has shown that airborne pollutants produced outside of the park boundaries have impacted the park's streams, vegetation, soils, and visibility. Sulfate, nitrates, and ozone are converted from sulfur dioxide and nitrogen oxides emitted from the burning of fossil fuels (Renfro 2001). Wind currents move the pollutants from industrial areas, urban areas, and automobiles into the southern Appalachian Mountains where they are trapped by the height and physical structure of the mountains. Since 1948, annual average visibility in the southern Appalachians has decreased by an average of 60%, 80% in the summer and 40% in the winter (National Park Service 2001).

Ground level ozone pollution is another disturbance factor linked to air pollution. Ninety plant species have shown injury and are directly threatened by increased ozone levels. Ozone exposure in the park is among the highest in the eastern United States. Ozone levels tend to be more severe with higher elevation, and along GSMNP ridgetops ozone levels are often twice that of nearby cities Knoxville and Atlanta.

GSMNP also receives some of the highest sulfur and nitrogen deposition rates in all monitored locations in North America. Between the years 1981 and 2000, annual wet nitrate deposition increased by 16% (Renfro 2001). Acid deposition causes base nutrient depletion, which is the leaching of nutrients from the soil. As the forest soils lose nutrients, the vegetation that relies on the soils for the nutrients become more sensitive to disease and stress. Nitrogen saturation causes toxic aluminum to be released, which further degrades the soil conditions and plant ecology. Along with the soils, streams have been acidified to the point that the park's high elevation ecosystems are in jeopardy. The heavy metals and aluminum released by acid deposition creates unsuitable living conditions for aquatic organisms. Some streams are approaching the public health standard limits for unacceptable nitrate levels.

Invasive species are another form of disturbance that concerns the park conservationists. Several non-native plants and animals reside in the park, many of which threaten the diversity of the park since they are not natural components of the ecological system. Kudzu (*Pueraria lobata* Willd.), mimosa (*Albizzia julibrissin* Durazz.), and Oriental bittersweet (*Celastrus orbiculatus* Thunb.) are a few examples of over 380 species of non-native plants found in GSMNP (NPS 2001). Many of these species occur in areas of historic logging and other disturbed vegetation areas where they outcompete native species.

The balsam woolly adelgid (*Adelges piceae* Ratzeburg) and the hemlock woolly adelgid are the most visible examples of invasive species in GSMNP. The balsam woolly adelgid is an insect that infests and destroys stands of Fraser firs found in the high elevations of the park. The adelgid infects the tree with toxins that block the path of nutrients to the tree. It is estimated that 95% of the park's mature Fraser firs have been attacked by the adelgid (Starkey 1997). The hemlock wooly adelgid moved into the park in May of 2002 and targets the hemlock stands that are commonly found in the park, especially along waterways. The insect slows or prevents tree growth by sucking sap from young twigs. This process causes the needles to discolor and prematurely drop which impairs the health of the tree (NPS 2003).

Wild hogs (*Sus scrofa* L.) are another example of a non-native species that threatens the natural ecologic communities in GSMNP. In 1912, as part of a business venture, the Whitting Manufacturing Company of England introduced the wild boar to establish a game preserve on Hooper Bald in Graham County, North Carolina where they were allowed to proliferate undisturbed for 8 to 10 years. By the 1920s, most of the wild boars had escaped from their crude enclosures into the surrounding terrain. Normal boar feeding behavior, called rooting, disturbs the soil. Hog rooting destroys vegetation, reduces small mammal habitat, and often leads to increased soil erosion (Keller *et al.* 2003).

Chapter Four
Methods and Procedures

4.1 Site Selection

The study sites were selected along the existing trail system and roadways of GSMNP. I began by first placing a cardboard plot sheet over a park map to identify potential sample areas. The sheet covered the entire park and systematically derived sections of the park to be sampled (Fig. 4.1). The trails and roadways that were contained by the delineated areas were identified and selected by trail or road name in ESRI ArcMap. The selected trails and roads were made into a layer that was overlaid with the environmental variable layers within the GIS to determine which combination of trails and roadways gave the best representation of the various geology types, soil types, vegetation types, logging types, slopes, aspects, and elevation ranges. Due to time constraints, not all delineated areas were visited, but a complete representation of the environmental factors was included in the sampling scheme. There were a total of 194 sample sites ranging from a minimum elevation of 354 m to a maximum elevation of 2003 m (Fig. 4.2).

4.2 Data Acquisition

From a predetermined starting point along a selected trail, sample points were established. A Garmin 12XL GPS or a Garmin eTrex GPS unit were used to measure 500 m increments along the trail to establish the sample points. If a GPS signal was not

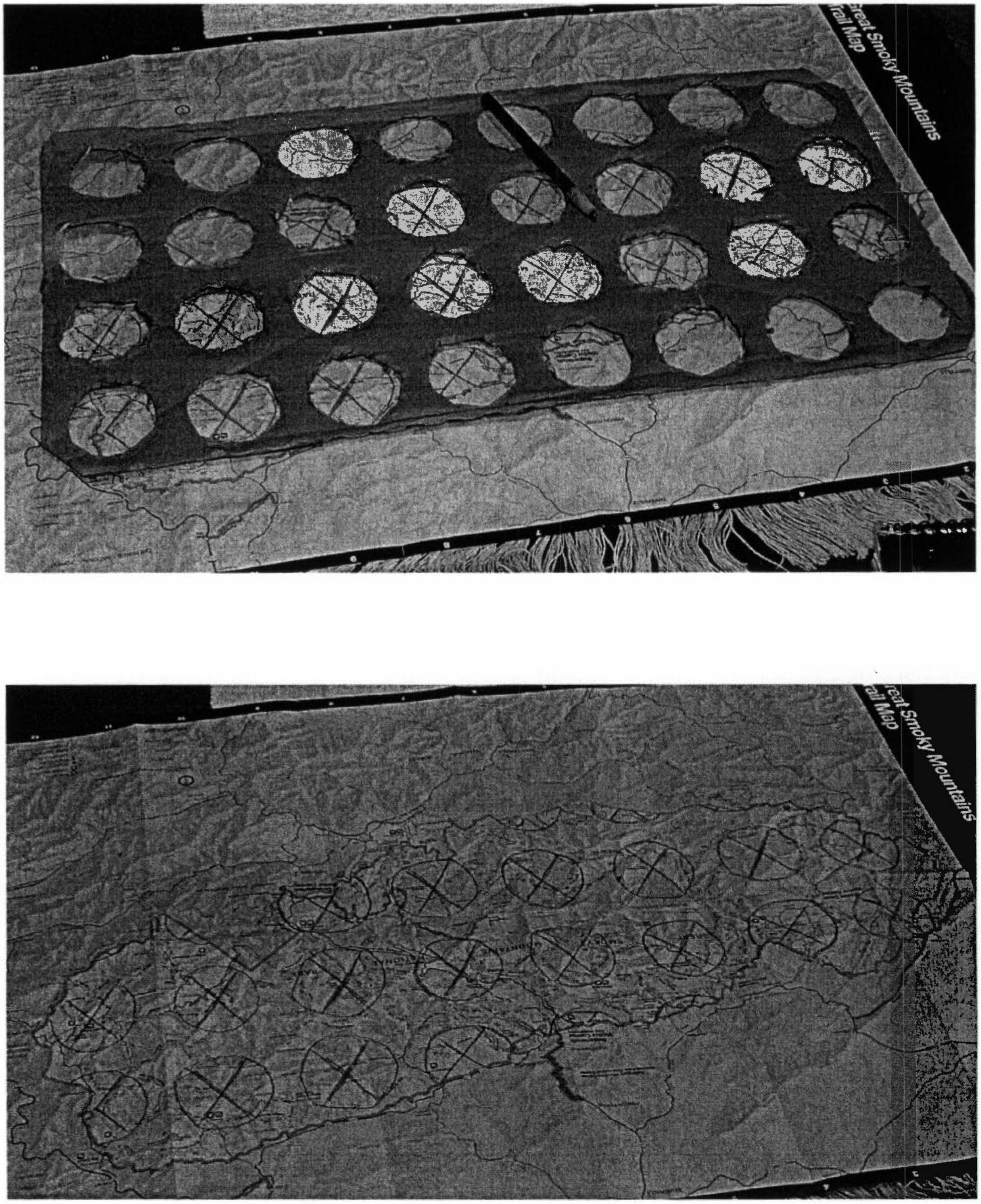


Figure 4.1. Systematic sampling grid. This method was used for study site delineation and to gain a complete coverage of the park and environmental variables.

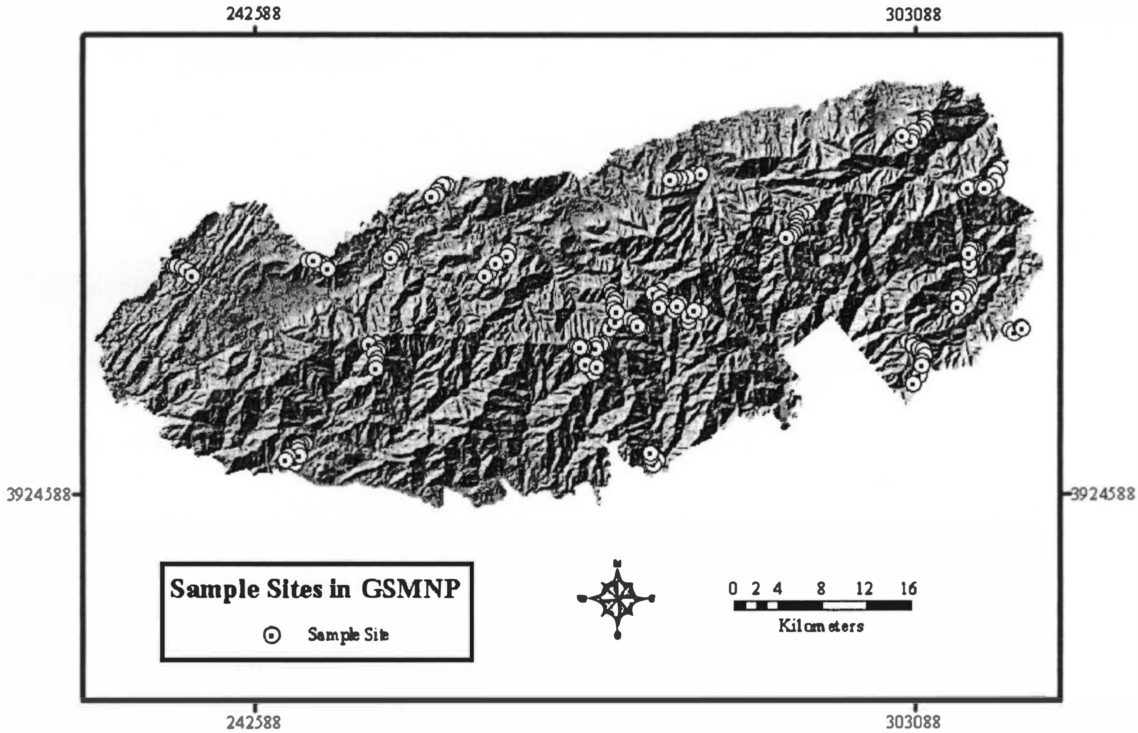


Fig. 4.2: Sample sites in GSMNP.

available or was blocked by the mountains or trees, pacing was used to measure the 500 m distance. On arrival at a sample point, a 15 m quadrat was created originating from the point on the trail. Using a pre-measured string, a 7.5 m radius was measured out in four directions from the center point and marked with flags to create the boundary of the quadrat. Once the sample site was laid out, a Trimble GeoExplorer III GPS unit was used to collect spatial information. A GPS point was recorded every five seconds and automatically averaged to determine the UTM coordinates of the site. This information was continuously gathered while the survey was being conducted, and at least 120 points were collected for each site. A Trimble GPS data dictionary was loaded into the unit for efficient on-site data entry (Table 4.1). Once the search was complete, site data were recorded into the GPS unit by way of the data dictionary and stored.

Each survey was conducted for ten minutes at each site. During each ten-minute count, the presence/absence of *V. latissimus* was recorded as well as the number of individuals that were present. The entire quadrat was sampled with the exception of steep slopes, extremely thick vegetation, or any other hazardous area. Close attention was given to more probable habitat features, such as beneath logs, moss, and stones, as well as under thick leaf litter and along the base of trees. Once the data were recorded in the GPS unit, the file was closed and another 500 m were measured along the trail to set up the next sample site. This procedure was followed for the remainder of the trail segment included in the study.

Table 4.1. Data dictionary fields. These fields were stored in the Trimble GeoExplorer III GPS unit for more efficient data entry in the field. Some fields had a menu option that allowed the user to choose from a preset list of inputs *i.e. latissimus*: Yes/No. Data were input using the GPS keypad for fields without the menu option.

Data Dictionary Fields
Date: date sample was taken
Time: time the sample site was visited
Trail: name of the trail being sampled
Site Number: site number along the trail
Latissimus: species present or absent
latissimus Count: number of individuals present
Weather: current weather conditions at the site
Temperature: current temperature at the site
Canopy Species: dominant canopy species in the quadrat
Pictures: were pictures taken (yes/no)
Comments: factors present that may be of importance

4.3 Data Organization

4.3.1 Geodatabase Structure

In this project, almost all the data were imported from outside sources and integrated with the project geodatabase. Much of the data came from the National Park Service (NPS), Natural Resource Conservation Service (NRCS), and the United States Geological Survey (USGS). Other data were converted from Trimble GPS rover files or created using ESRI ArcGIS Spatial Analysis. Vector data were imported and converted from shapefiles into feature classes. This process stores the geometry and attribute information and defines the coordinate system to match the coordinate system of the feature dataset (MacDonald 2001). All vector data were stored and managed through the personal geodatabase, and raster data were kept in an adjacent folder outside the geodatabase (Fig. 4.3). The data created from the geodatabase source data, such as hillshades, slope, and aspect, were automatically created in the source destination folder with the rest of the raster data. This organization structure allowed all the data to be managed together.

4.3.2 Projection Information

The coordinate system for this project was defined in the North American Datum of 1927 (NAD 1927), which is based on the Clarke Spheroid of 1866. A Geodetic datum anchors the coordinate reference system to a model that describes the size and shape of the earth. The NAD 1927 is the datum used by GSMNP GIS personnel and was used in this project to keep the data standard for future NPS implementation. The coordinate

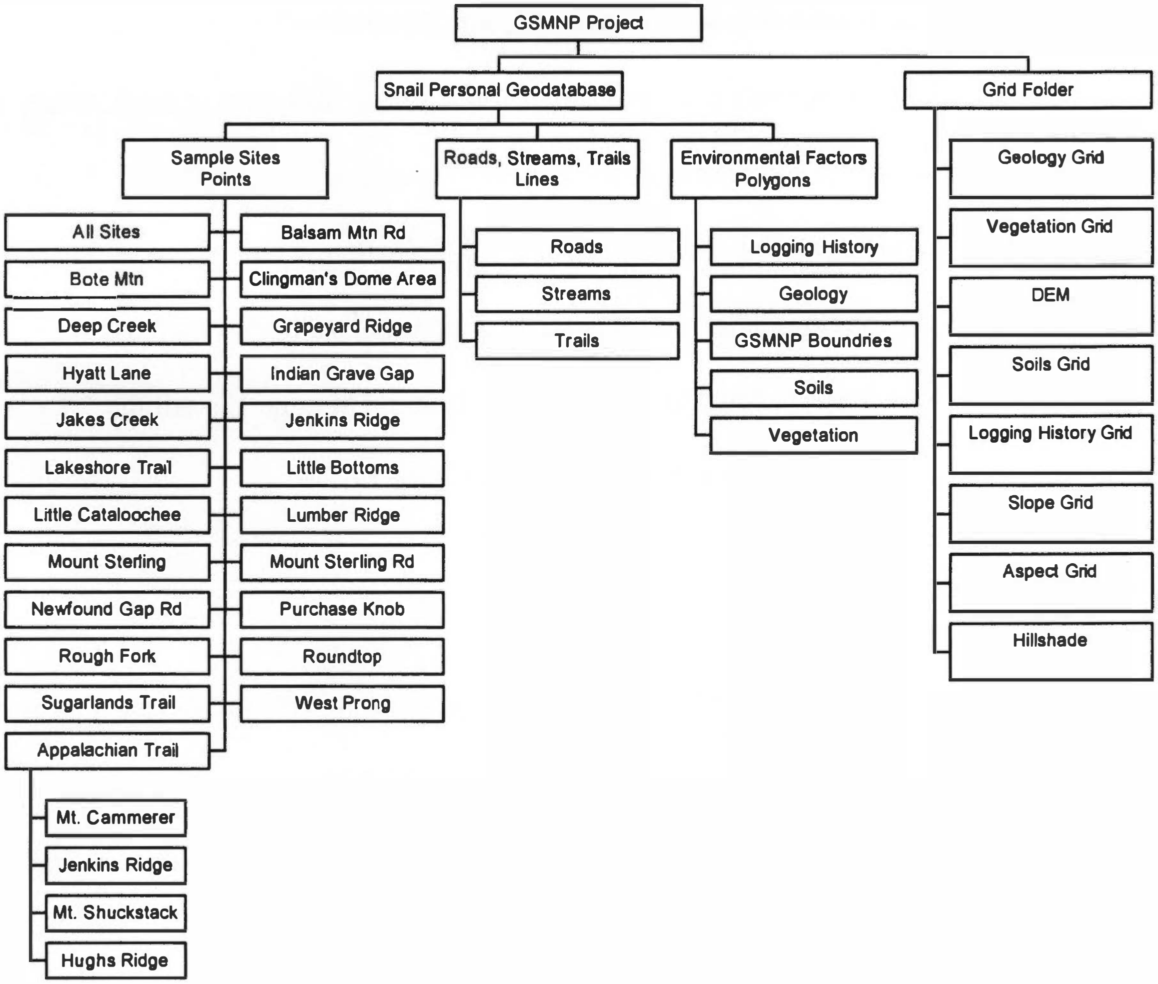


Fig. 4.3. Flow chart of GIS data organization. Data were stored in a common folder that contained the personal geodatabase as well as the grid folder that managed the environmental data.

system used within the NAD 1927 is the Universal Transverse Mercator (UTM), a projected coordinate system that divides the world into 60 north and south zones, six degrees wide. The northing values are measured continuously from zero at the Equator, in a northerly direction. UTM map units are measured in meters. The UTM coordinate system is based on the transverse Mercator conformal projection, which preserves shapes and angles along a central meridian and is good for relatively small areas such as GSMNP. The Great Smoky Mountains are located within UTM zone 17. All the following data were projected using NAD 1927 UTM Zone 17.

4.3.3 Vector Data

The point data in this project were managed in the "Sample Sites" feature dataset (Table 4.2). All 194 sample sites surveyed for the project were represented as point features on the map and were located with a single coordinate-pair value. A feature class containing the sample sites located on a particular stretch of trail was used to represent each trail surveyed. The data were collected using a Trimble GeoExplorer III GPS unit and were imported from GPS rover files. The shapefiles imported to the geodatabase contained all the attribute information that was collected in the field. The attributes included date, time, weather, canopy species, species present/absent, number of individuals, comments, and the UTM coordinates.

Once all the data were collected, the sample sites were merged together to create an inclusive feature class that contained attribute and spatial information for all 194 sites. Environmental factor attributes were spatially joined and entered for each discrete site.

Table 4.2. GIS data files used for the project

File Name	File Type	Description	Source
Sample Sites	Vector Point	Joined GPS sample points	GPS field data
Logging History	Vector Polygon	Logging history of GSMNP	NPS
Geology	Vector Polygon	Geology of GSMNP	NPS
GSMNP	Vector Polygon	GSMNP boundaries	NPS
Roads	Vector Line	Roads of GSMNP	NPS
Soils	Vector Polygon	Soils of GSMNP	USGS
Streams	Vector Line	Streams of GSMNP	NPS
Trails	Vector Line	Trails of GSMNP	NPS
Vegetation	Vector Polygon	Forest cover of GSMNP	NPS
elev_gsmnp	DEM	Elevation of GSMNP	USGS
Vegetation	GRID	Grid layer of vegetation	NPS
Geology	GRID	Grid layer of geology	NPS
Soils	GRID	Grid layer of soils	Spatial Analyst conversion
Slope	GRID	Slope values of GSMNP	Spatial Anaylst calculation
Aspect	GRID	Aspect values of GSMNP	Spatial Analyst calculation
Hillshade	GRID	Surface illumination of GSMNP	Spatial Analyst calculation

Each point in the combined feature class contained all the field attributes as well as environmental values such as geology, soils, slope, aspect, elevation, vegetation, and logging history.

The linear features in this project were used to show relative location of the sample sites during the data collection and analysis phase. The linear features included roads, streams, and trails (Table 4.2). These features were overlaid with environmental factors to determine which trails and roads should be sampled so that all classes within each environmental factor would be represented in the model. Each feature class contained attribute information, which included road, stream, or trail name, as well as the length of each particular line segment. The roads, trails, and streams data were created by the NPS by digitizing existing USGS quads. Trails that did not appear on the USGS maps were created with a GPS. The data were obtained from NPS personnel located at the Twin Creeks Natural Resource Center in Gatlinburg, Tennessee. The data are managed in the "Roads Streams and Trails" feature dataset within the geodatabase.

The polygon data in this project represented the environmental variables that could be factors in the spatial distribution of land snails and other organisms. Polygon data included geology, soils, overstory vegetation, and logging history (Table 4.2). The polygon data were spatially joined to the combined feature class for further statistical analysis. The data were also used to help identify which trails and roads should be sampled to best cover the range of environmental data. Each polygon feature class contained attribute information including names and unit codes of each class within the feature and other descriptive information about the data.

Some of the data, such as geology and logging history, were obtained through NPS personnel at Twin Creeks Natural Resource Center. Soils data were obtained from USGS personnel and data (such as vegetation) were converted from raster files obtained from the Park Service. The data are stored in the "Environmental Factors" feature dataset.

4.3.4 Grid Data

The raster data in this project were used in the calculations for the model. Raster datasets are best for representing continuous features with gradually varying attributes. Raster data represented the environmental factors that were found to be statistically significant. Both spatially continuous data and spatially discrete data were used for the analysis in this project. Spatially continuous data represent measured quantities and included elevation, slope, and aspect. Spatially discrete data represent categorical or classified data including vegetation, geology, soils, and logging history (Table 4.2). Slope and aspect were derived from the Digital Elevation Model (DEM), which was used to represent elevation. The Spatial Analyst extension in ESRI ArcGIS was used to perform this terrain analysis as well as to generate a hillshade model to simulate the illumination of the surface based on sun angle and slope of the surface. The hillshade was used to offer better visual representation of elevation.

Elevation, slope, and aspect are continuous data and were not converted to polygons to be spatially joined to the combined feature class due to the size of the files. The values for these factors were overlaid with the combined feature class and elevation, slope, and aspect values were obtained by using the "identify" tool in ESRI ArcGIS. The values were then manually input into the combined database to be used in the statistical analysis.

The vegetation coverage was originally created using two adjacent Landsat-5 TM scenes and an unsupervised classification technique yielding 14 classes. An overall accuracy of 83% was obtained using 271 field plots. The mapping was conducted by the Environmental Remote Sensing Center, Institute for Environmental Studies, University of Wisconsin-Madison. The data are dynamic, and the accuracy has the possibility of changing over time. Park personnel digitized existing USGS geology maps to create the generalized geology maps for the park. Raster data obtained from the NPS included the DEM, vegetation, geology, and logging history. The soils layer was converted from a polygon shapefile obtained from the USGS. The raster data are found within the geodatabase in the "Grids" folder.

4.4 Data Preparation

Before the statistical analysis could be conducted or the habitat model constructed, the field data had to be corrected and transferred from the GPS unit to a useful GIS format, which was completed through the GPS Pathfinder Office interface developed by Trimble. Due to the nature of GPS, many possible sources for error can exist in the datafiles. The signal sent from the satellite to the GPS receiver passes through the ionosphere and the interaction with the charged particles slows the signal. Multipathing errors occur when the signal does not come directly to the receiver but bounces off buildings, or in this case, mountains. A small amount of error is associated with the satellites and receivers themselves. To account for this error, the data are differentially corrected within Pathfinder Office before they are exported to the GIS. Files are downloaded from a GPS base station in Asheville, North Carolina that represent the correct timing information for each satellite during a particular hour of a particular day. The downloaded base station files are compared to the rover files collected in the field and automatically adjusted within Pathfinder's differential correction utility. The error in the sample site data has been corrected and the data are now accurate to an average of +/– 1 m.

For the sample site data to be used in the GIS, they had to be transferred to a recognizable GIS format. The "Export" utility was used to complete this task. The corrected datafiles were selected for transfer and in addition to the data dictionary attributes, additional attributes were selected to be included in the attribute table. These attributes were correction status, receiver type, position, height, and position dilution of precision (PDOP). The selected files and attributes were exported from the GPS format to ArcView shapefiles. The files were then ready to be imported into the project geodatabase, and this was accomplished by using the "shapefile to geodatabase" tool in ArcCatalog. The GPS derived shapefiles were imported into the correct feature dataset within the project geodatabase and the data transfer was complete. All the attributes collected in the field on the Trimble GeoExplorer III GPS unit have now been integrated into the GIS platform and are ready for analysis.

This process was conducted after every field session. The data were stored in the geodatabase as individual trail point files until all sites had been sampled. Once all the sites had been transferred to the GIS format and imported to the project geodatabase, they were merged into one point file containing all 194 sites and attributes named "all_sites" (Appendix B). The next step in preparing the data was to spatially join the environmental factor data to each sample site. Each environmental factor that was represented in polygon form (geology, vegetation, soil, logging history) was spatially joined to the sample sites (Appendix B). All the attribute information for each environmental factor was listed for each sample site in the "all_sites" data table, and the geology, vegetation, soil, and logging history for any given sample site was determined.

A different procedure was used to include the continuous environmental factors elevation, slope, and aspect. These three environmental factors were layered on top of one another in raster format. The "all_sites" point file was added. Using the identify tool, the value for each factor was determined at each site. In ArcCatalog, new data fields for each factor were created in the "all_sites" data table and populated with the value data obtained for each site (Appendix B).

4.5 Statistical Analysis

Statistical analysis was conducted to identify relationships between the seven environmental variables (vegetation, soils, geology, logging history, elevation, slope, and aspect) and the presence or absence of *V. latissimus* at each sample site. The SPSS statistical software package was used to perform the analysis. The analysis was conducted to determine which variables influence the distribution of the snail and therefore should be included in the habitat suitability model.

First, the data were exported from the geodatabase to an SPSS file format. Categorical data with more than five classes were reclassified before conducting analysis due to the relatively small sample size for an area as large as GSMNP. This assured that there would be multiple sites for each class for more valid results. Each factor was reclassified into five categories based on similar characteristics.

A χ^2 analysis was used to assess whether possible relationships existed between the presence/absence data and the environmental factors. This analysis is used for testing relationships between two categorical variables such as the variables in this study. The Crosstabs command in SPSS was used to evaluate possible relationships present in the data, as well as perform the χ^2 analysis. Crosstabs is an SPSS procedure that cross-tabulates two variables, and displays their relationship in tabular form and generates information about bivariate relationships. The test reports an adjusted residual value. An adjusted residual greater than two shows that the class is significant to the presence of the species. If the adjusted residual is between two and negative two, the class shows little difference between "Yes" occurrences and "No" occurrences, and an adjusted residual less than –2 shows that the class is significant to the absence of the species. Crosstabs also creates a table that contains a cell for every combination of categories in the two variables with the number of cases meeting the particular combination as well as the percentage of cases of all the cells. This aspect makes the test useful for breaking down percentages within each class of a variable.

The dependent variable was the presence/absence data, and the independent variables were the environmental factors. Parameters were set to report the observed counts, the row, column, and total percentages, and the adjusted residuals for each class. Pearson's χ^2 was used to indicate the strength of the relationships present. The results of these analyses report the degree to which the conditional distributions differ from what would be expected under the assumption of statistical independence. The null hypothesis states that no relationship exists between the variables in the bivariate table, *i.e.,* the variables are statistically independent. A rejection of the null hypothesis suggests a possible relationship exists between the environmental variables and the presence/absence data.

Continuous data such as elevation, slope, and aspect could not be analyzed using χ^2 and crosstabs procedures. An independent two-sample t-test was used in SPSS to identify the level of significance between these continuous variables and the presence of *V. latissimus.* This test is good to use when the population variance is unknown and the sample size is small. The test compares the two sample means to test for significance between the two variables. The 95% confidence interval of the difference was used to determine significance. The SPSS process reported the t value, degrees of freedom, the two-tailed significance level, and the mean difference.

Logistic regression was used to assign weights to factors that were found to be significant with the presence of *V. latissimus.* In SPSS, binary logistic regression was performed with the presence/absence data as the dependent variables and a significant environmental factor as the independent variable. The test affirmed the significance of the variable and determined the strength of the relationship. Weights were assigned based on strength of relationships and the percentage of the data that was correctly classified by the logistic regression. Weights were assigned as decimals with a sum equal to one so that the resulting suitability values would range between one and five.

4.6 Data Processing

Once the statistical analysis was conducted to identify relationships between environmental factors and the presence of *V. latissimus,* the data were reclassified based upon these relationships. The ArcGIS Spatial Analyst command "Reclassify Grid" was used to perform the reclassification of each significant grid layer (Appendix B). Each class within a variable was given a new value based on its ranking. Pixel values of five represented the optimal *V. latissimus* range for each environmental factor down to a pixel value of one that represents the least suitable range for the species within the factor. Once all significant layers had been reclassified, the data were ready to be analyzed for the optimal *V. latissimus* habitat suitability areas in GSMNP.

The ArcGIS Spatial Analyst tool "Raster Calculator" was used to construct the habitat suitability model (Appendix B). Significant factors were input into the equation with weights that were established based upon the significance of each factor to the presence of *V. latissimus* derived from the logistic regression analysis. The suitability zone five represented the optimal habitat area for *V. latissimus* in the park, and the suitability zone one represented the least suitable habitat area in the park for the species.

4.7 Species Abundance Model

A species abundance model was constructed to compare the populations of a variety of land snails in the park. This analysis investigated two species of snails in the Clingman's Dome area of the park. The abundance of the primary species of this study, *Vitrinizonites latissimus,* was compared to the abundance of *Stenotrema depilatum* to make further observations about land snail distributions in GSMNP.

The sample sites feature class was added twice to the data frame to display the abundance of both snails. Under layer properties, the symbology was changed from single symbol to quantities. The proportional symbols option was used to map the abundance of each species. Proportional symbols represent data values more precisely as the size of a proportional symbol reflects the actual data value. The number of individuals of each species recorded at each site was used as the data value. The data exclusion option of the layer properties was used to remove the zero values from the data for each species because zero values had to be excluded from the procedure so that proportions could be mapped correctly. The *V. latissimus* data were broken into five classes ranging from one snail to ten snails at a site, and *S. depilatum* was represented by one class representing two snails at a site. The significant environmental layers used for the habitat suitability model were added to the data frame and abundance measurements were compared to environmental factors to evaluate whether any relationships could be seen between the number of species present at a site and a particular class of an environmental factor.

Chapter Five
Results

5.1 Cross-tabulation and Chi-Squared Results

Five reclassified geology classes (Table 5.1) were analyzed using the SPSS crosstabs procedure (Table 5.2) and Pearson's χ^2 test for significance. Class one was the only significant geology class with 71.8% of the count having positive occurrences and a "Yes" adjusted residual of 5.1. Class two was the next most significant geology class with 61.4% of the count having positive occurrences and a "Yes" adjusted residual of 1.5. Class three was the third most significant geology class with 75.0% of the count having positive occurrences and a "Yes" adjusted residual of 0.9. Class five had the least significant values with 100.0% of the count having negative occurrences and a "No" adjusted residual of 3.7. Class four was the second least significant geology class with only 10.3% of the count having positive occurrences and a "Yes" adjusted residual of -0.8. The χ^2 value for geology was 55.883 (df = 4, $P < 0.0001$).

Five reclassified vegetation classes (Table 5.3) were analyzed using the SPSS crosstabs procedure (Table 5.4) and Pearson's χ^2 test for significance. Class one was the most significant vegetation class with a high "Yes" adjusted residual of 8.7. There were positive occurrences of the snail at 95.3% of the sites sampled within this class. Class three and class two were the next most significant vegetation classes. Class three had a "Yes" adjusted residual of 0.7 with 60.0 % of the count having positive occurrences. Fifty percent of the class two count had positive occurrences with a "Yes" adjusted residual of -0.2. Class five and class four had the least significant values for the

Table 5.1. Reclassification values for geology

Reclassified Value	Original Geology Values
Class One	Thunderhead Sandstone
Class Two	Anakeesta Formation, Great Smoky Group
Class Three	Rich Butt Sandstone
Class Four	Roaring Fork Sandstone, Elkmont Sandstone, Cades Sandstone, Longarm-Quartzite, Basement Complex, Wilhite Formation Coarse
Class Five	Metcalf Phyllite, Wilhite Formation

Table 5.2. Cross-tabulation results for geology

Reclassified Geology Classes		***V. latissimus* Present/Absent**		**Total**
		No	Yes	
1	Count	24	61	85
	% within Class 1	28.2%	71.8%	100.0%
	Overall %	27.0%	64.2%	46.2%
	Adjusted Residual	-5.1	5.1	
2	Count	17	27	44
	% within Class 2	38.6%	61.4%	100.0%
	Overall %	19.1%	28.4%	23.9%
	Adjusted Residual	-1.5	1.5	
3	Count	1	3	4
	% within Class 3	25.0%	75.0%	100.0%
	Overall %	1.1%	3.2%	2.2%
	Adjusted Residual	-0.9	0.9	
4	Count	35	4	39
	% within Class 4	89.7%	10.3%	100.0%
	Overall %	39.3%	4.2%	21.2%
	Adjusted Residual	5.8	-5.8	
5	Count	12	0	12
	% within Class 5	100.0%	0%	100.0%
	Overall %	13.5%	0%	6.5%
	Adjusted Residual	3.7	-3.7	
Total	Count	89	95	184
	% within Geology	48.4%	51.6%	100.0%
	Overall %	100.0%	100.0%	100.0%

Table 5.3. Reclassification values for vegetation

Reclassified Value	Original Vegetation Values
Class One	Northern Hardwood, Spruce-Fir
Class Two	Cove Hardwood, Grassy Bald
Class Three	Mesic Oak
Class Four	Mixed Mesic Hardwood, Treeless
Class Five	Pine, Xeric Oak, Tulip Poplar, Pine-Oak

Table 5.4. Cross-tabulation results for vegetation

Reclassified Vegetation Classes		*V. latissimus* Present/Absent		Total
		No	Yes	
1	Count	3	61	64
	% within Class 1	4.7%	95.3%	100.0%
	Overall %	3.3%	64.2%	34.4%
	Adjusted Residual	-8.7	8.7	
2	Count	23	23	46
	% within Class 2	50.0%	50.0%	100.0%
	Overall %	25.3%	24.2%	24.7%
	Adjusted Residual	0.2	-0.2	
3	Count	6	9	15
	% within Class 3	40.0%	60.0%	100.0%
	Overall %	6.6%	9.5%	8.1%
	Adjusted Residual	-0.7	0.7	
4	Count	29	2	31
	% within Class 4	93.5%	6.5%	100.0%
	Overall %	31.9%	2.1%	16.7%
	Adjusted Residual	5.4	-5.4	
5	Count	30	0	30
	% within Class 5	100.0%	0%	100.0%
	Overall %	33.0%	0%	16.1%
	Adjusted Residual	6.1	-6.1	
Total	Count	91	95	186
	% within Vegetation	48.9%	51.1%	100.0%
	Overall %	100.0%	100.0%	100.0%

vegetation variable, class five with 100.0% of the count having negative occurrences and class four with only 6.5% of the count having positive occurrences. Class five had a "No" adjusted residual of 6.1 and class four had a "Yes" adjusted residual of -5.4. The results of the χ^2 analysis showed that vegetation was a significant variable ($\chi^2 = 106.642$, df = 4, P < 0.0001).

The SPSS Crosstabs procedure and Pearson's χ^2 test for significance were used to determine the significance of the relationship between five reclassified soil classes (Table 5.5) and the presence or absence of *V. latissimus*. Crosstabs showed significant results (Table 5.6) for the relationship. The most significant soil class, class one, had 89.6% of the count with positive occurrences and a "Yes" adjusted residual of 6.0. Class two also showed significant results with 84.4% of the count having positive occurrences and a "Yes" adjusted residual of 5.0. The remaining classes were not significant factors for the presence of the snail. Class three had a large count, but only 28.1% of the count had positive occurrences. The "Yes" adjusted residual for class three was -4.3. Class four had a smaller count and a "Yes" adjusted residual of -1.6. Twenty-five percent of the count had positive occurrences in class four. Class five did not have any positive occurrences and had a "No" adjusted residual of 6.5. The Pearson's χ^2 test supported the significant results for soils. The χ^2 value was 106.642 (df = 4, P < 0.0001).

There were only five logging history classes (Table 5.7) and the variable did not need to be reclassified to perform the SPSS crosstabs procedure or the Pearson's χ^2 test for significance. The crosstabs results (Table 5.8) did not show a strong relationship between logging history and the presence of *V. latissimus*. The only significant logging

Table 5.5. Reclassification values for soils

Reclassified Value	Original Soil Units
Class One	Breakneck-Pullback
Class Two	Oconaluftee-Guyot-Chiltoskie, Luftee-Anakeesta
Class Three	Wayah-Tanasee, Soco-Stecoah-Spivey, Dellwood-Smokemont-reddies
Class Four	Cataska-Sylco-Spivey
Class Five	Junaluska-Tsali, Ditney-Unicoi-Spivey, Lauda-Fanning, Evard-Cowee-Leatherwood

Table 5.6. Cross-tabulation results for soils

Reclassified Soil Classes		*V. latissimus* Present/Absent		Total
		No	Yes	
1	Count	5	43	48
	% within Class 1	10.4%	89.6%	100.0%
	Overall %	5.5%	43.4%	25.3%
	Adjusted Residual	-6.0	6.0	
2	Count	7	38	45
	% within Class 2	15.6%	84.4%	100.0%
	Overall %	7.7%	38.4%	23.7%
	Adjusted Residual	-5.0	5.0	
3	Count	41	16	57
	% within Class 3	71.9%	28.1%	100.0%
	Overall %	45.1%	16.2%	30.0%
	Adjusted Residual	4.3	-4.3	
4	Count	6	2	8
	% within Class 4	75.0%	25.0%	100.0%
	Overall %	6.6%	2.0%	4.2%
	Adjusted Residual	1.6	-1.6	
5	Count	32	0	32
	% within Class 5	100.0%	0%	100.0%
	Overall %	35.2%	0%	16.8%
	Adjusted Residual	6.5	-6.5	
Total	Count	91	99	190
	% within Soils	47.9%	52.1%	100.0%
	Overall %	100.0%	100.0%	100.0%

Table 5.7. Class values for logging history

Class Value	Original Values
Class One	Selective Cut
Class Two	Light Cut
Class Three	Heavy Cut
Class Four	Undisturbed
Class Five	Settlement

Table 5.8. Cross-tabulation results for logging history

Logging History Classes		*V. latissimus* Present/Absent		Total
		No	Yes	
1	Count	27	11	38
	% within Class 1	71.1%	28.9%	100.0%
	Overall %	30.3%	11.6%	20.7%
	% of Total	14.7%	6.0%	20.7%
	Adjusted Residual	3.1	-3.1	
2	Count	19	9	28
	% within Class 2	67.9%	32.1%	100.0%
	Overall %	21.3%	9.5%	15.2%
	% of Total	10.3%	4.9%	15.2%
	Adjusted Residual	2.2	-2.2	
3	Count	14	20	34
	% within Class 3	41.2%	58.8%	100.0%
	Overall %	15.7%	21.1%	18.5%
	% of Total	7.6%	10.9%	18.5%
	Adjusted Residual	-0.9	0.9	
4	Count	8	52	60
	% within Class 4	13.3%	86.7%	100.0%
	Overall %	9.0%	54.7%	32.6%
	% of Total	4.3%	28.3%	32.6%
	Adjusted Residual	-6.6	6.6	
5	Count	21	3	24
	% within Class 5	87.5%	12.5%	100.0%
	Overall %	23.6%	3.2%	13.0%
	% of Total	11.4%	1.6%	13.0%
	Adjusted Residual	4.1	-4.1	
Total	Count	89	95	184
	% within Logging History	48.4%	51.6%	100.0%
	Overall %	100.0%	100.0%	100.0%
	% of Total	48.4%	51.6%	100.0%

history class was class four. The majority of the data, 60 sites, occurred in this class, and 86.7% of the class four count had positive occurrences with a "Yes" adjusted residual of 6.6. No other classes were significant. Class three showed 58.8% of the count with positive occurrences and a "Yes" adjusted residual of 0.9. Classes one, two, and five all had the majority of the count having negative occurrences. Class one had a "No" adjusted residual of 3.1, class two had a "No" adjusted residual of -2.2, and Class five had a "No" adjusted residual of -4.1. The Pearson's χ^2 test, however, showed that a significant association may be present between the logging history of the park and the presence of *V. latissimus*. The variable was significant ($P < 0.0001$) with a χ^2 value of 56.999 (df = 4).

Slope, aspect, and elevation could not be analyzed using crosstabs due to the continuous nature of these data. An independent sample t- test was used to test for the significance of these variables (Table 5.9). Elevation was significant ($t = 12.851$, $P < 0.0001$), while slope was not significant ($t = 1.879$, $P < 0.062$). Aspect also was not significant ($t = -0.979$, $P < 0.329$).

5.2 Assessing Variable Weight Assignment Using Logistic Regression

Logistic regression was conducted for each significant variable and the presence/absence of *V. latissimus* to verify the results from the frequency analysis and to assign a weight to the variable. Regression analysis conducted for elevation correctly classified 87.4 % of the data ($P < 0.0001$). Elevation was the most significant variable included in the model and received a weight of 0.50. Regression analysis conducted for

Table 5.9. T-test results for slope, aspect, and elevation

Variable	t-score	df	Significance (2-tailed)	Mean Difference
Slope	1.879	192	$P < 0.062$	2.466
Aspect	-0.979	192	$P < 0.329$	-14.56
Elevation	12.851	192	$P < 0.0001$	670.622

soils correctly classified 86.3% of the data ($P < 0.0001$). The soil variable received a weight of 0.30. Logistic regression was conducted for vegetation and 79.5% of the data were correctly classified ($P < 0.0001$). When compared to the other variables to be used in the model, vegetation was the least significant variable. Vegetation received a weight of 0.20. Regression conducted for geology correctly classified 52.1% but was not significant ($P < 0.477$), and therefore was not included in the habitat suitability model. The logging history variable correctly classified 64.7% of the data in the regression model ($P < 0.55$). The logging history variable was therefore not included in the model.

5.3 Variable Reclassification Results

The result of the vegetation reclassification was a new grid layer with pixel values ranging from one to five (Fig. 5.1). Vegetation class one was given a rank of five for the grid reclassification procedure and model generation as it showed the strongest relationship with the presence of the land snail. Class three was the next most significant vegetation class and was given a rank of four, class two, the third most significant class, was given a rank of three. Class four was the second least significant class and was given a rank of two for the grid reclassification procedure and model generation. It would be highly unlikely to find the snail in this vegetation class. Class five had the least significant values and was given a rank of one for the grid reclassification procedure and model generation. Snails are least likely to occur in this vegetation class.

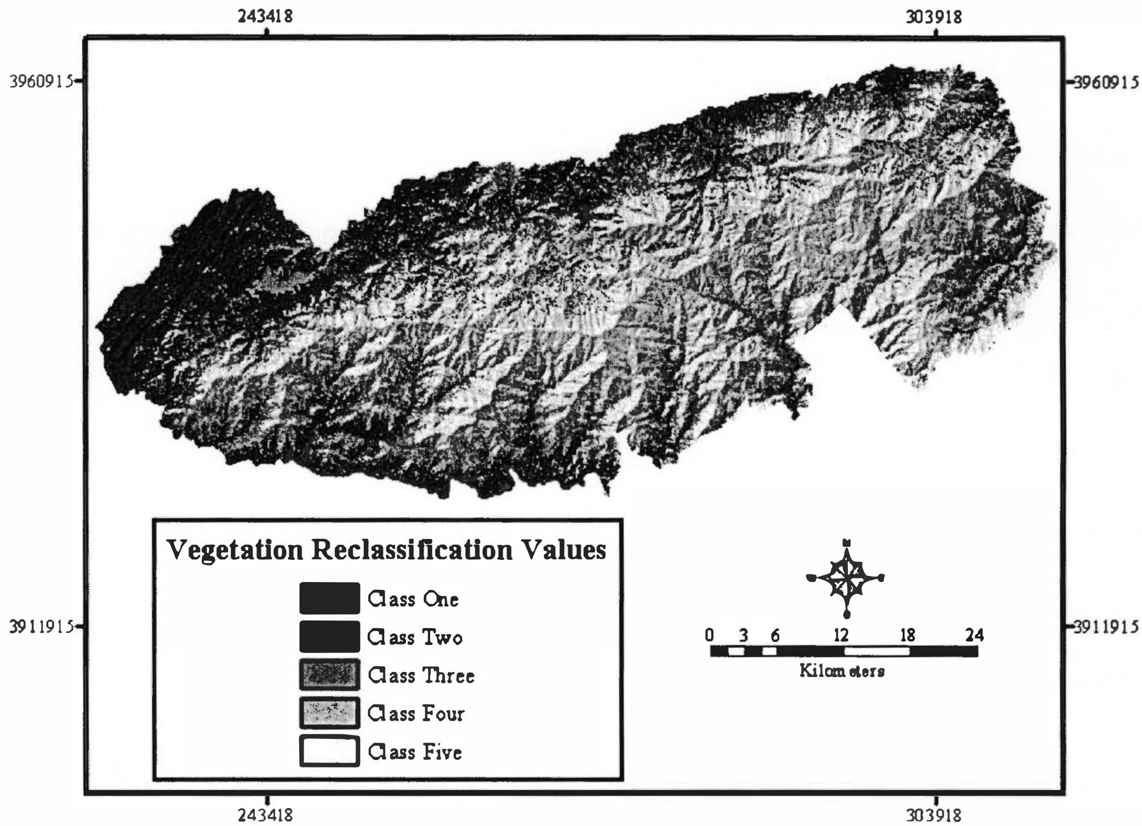


Fig. 5.1. Reclassified vegetation layer.

Based on the results of the statistical analysis conducted on the soils variable including Pearson's χ^2 test for significance, cross-tabulation, and logistic regression, the variable was reclassified (Fig. 5.2) and included in the habitat suitability model. Class one and class two showed the strongest associations with the presence of *V. latissimus*. The snail should be found in relative abundance in these zones. Class one was given a rank of five for the model generation as it was the most significant class and class two was given a rank of four. Classes three and four were similar in their results. Class three was the third least significant soil class, but it had more positive occurrences than class four and therefore was given a rank of three for the habitat model generation and grid reclassification procedure. Class four was given a rank of two and class five was given a rank of one for the grid reclassification procedure and model generation. It would be unlikely to find the snail in these zones.

Finally, elevation was the only continuous variable to be included in the habitat suitability model. Based on the t-test scores, slope and aspect were not found to be significant. The DEM was reclassified (Fig. 5.3) based on histograms of presence/absence data and elevation (Fig. 5.4). Increasing elevations contain greater abundance of the snail. Elevations ranging between 0–900 m were given a rank of one, as it is the least suitable habitat range for *V. latissimus*. Elevations ranging between 901–1100 m were given a rank of two, elevations between 1101–1400 m were given a rank of three, while elevations above 1801 m were given a rank of four. The optimal elevation range for this species lies between 1401–1800 m and was given a rank of five.

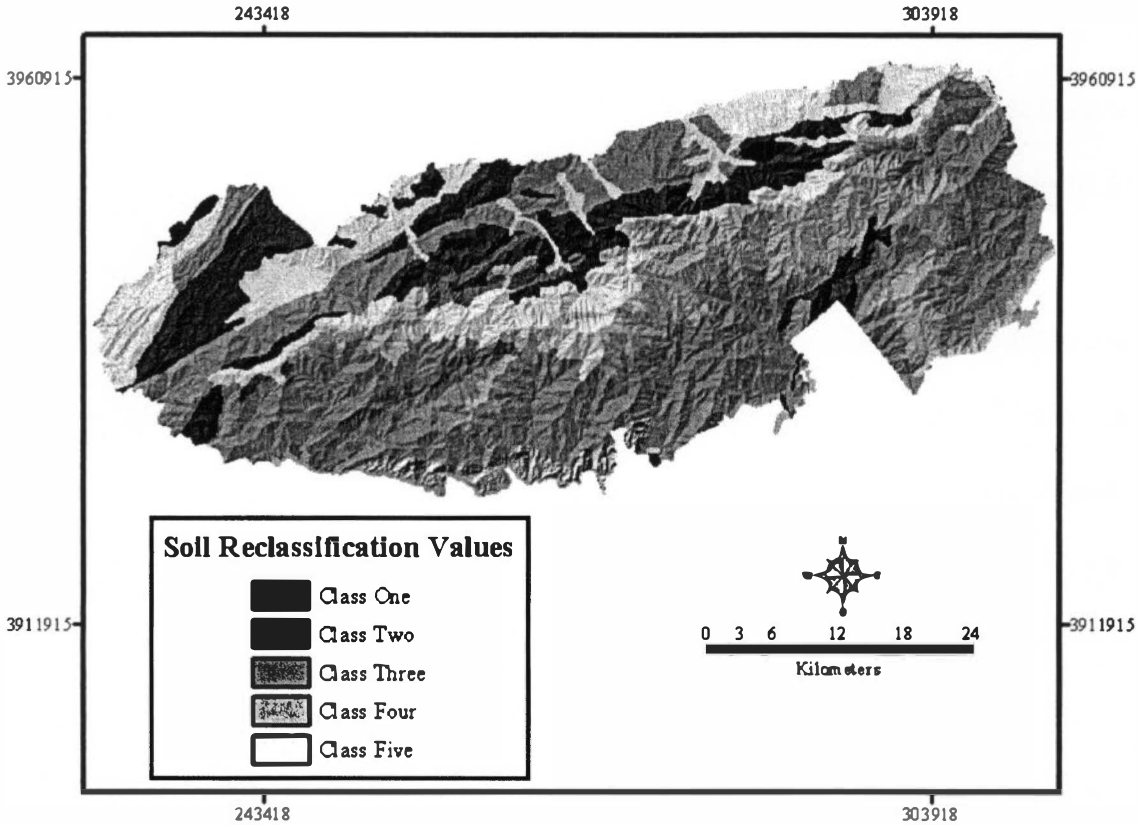


Fig. 5.2. Reclassified soil layer.

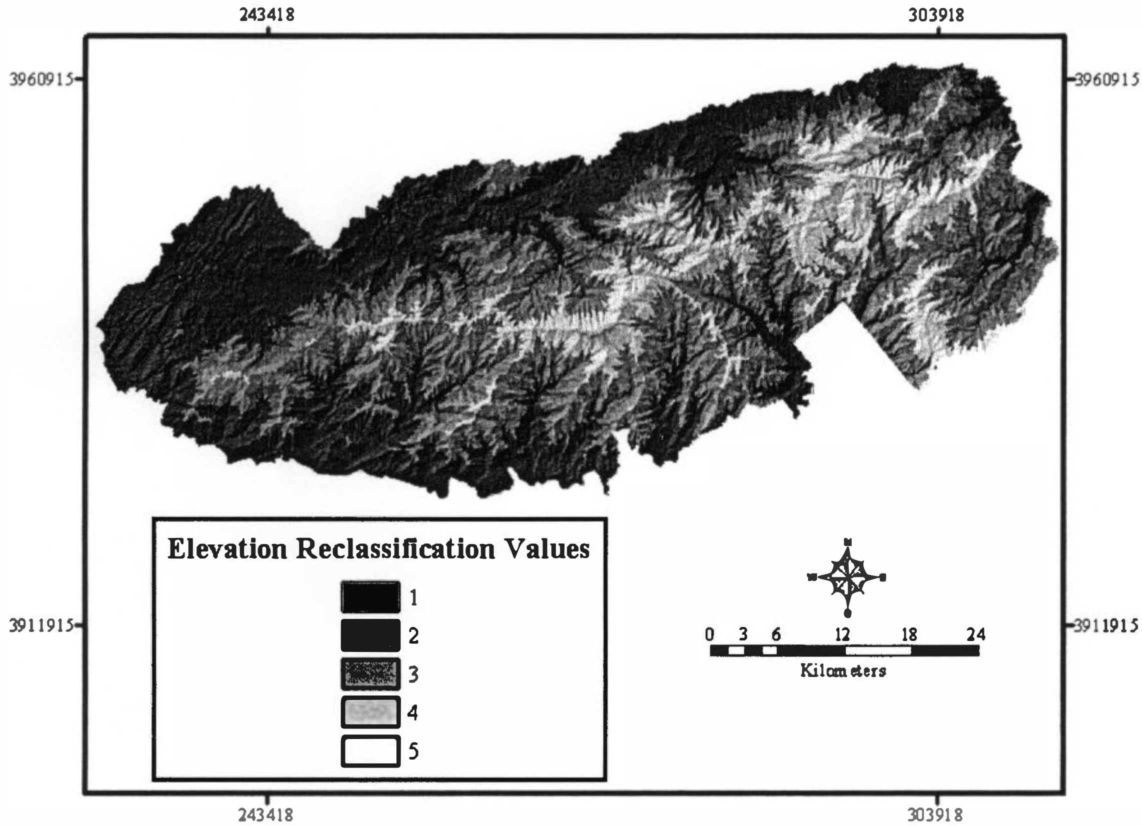


Fig. 5.3. Reclassified elevation layer.

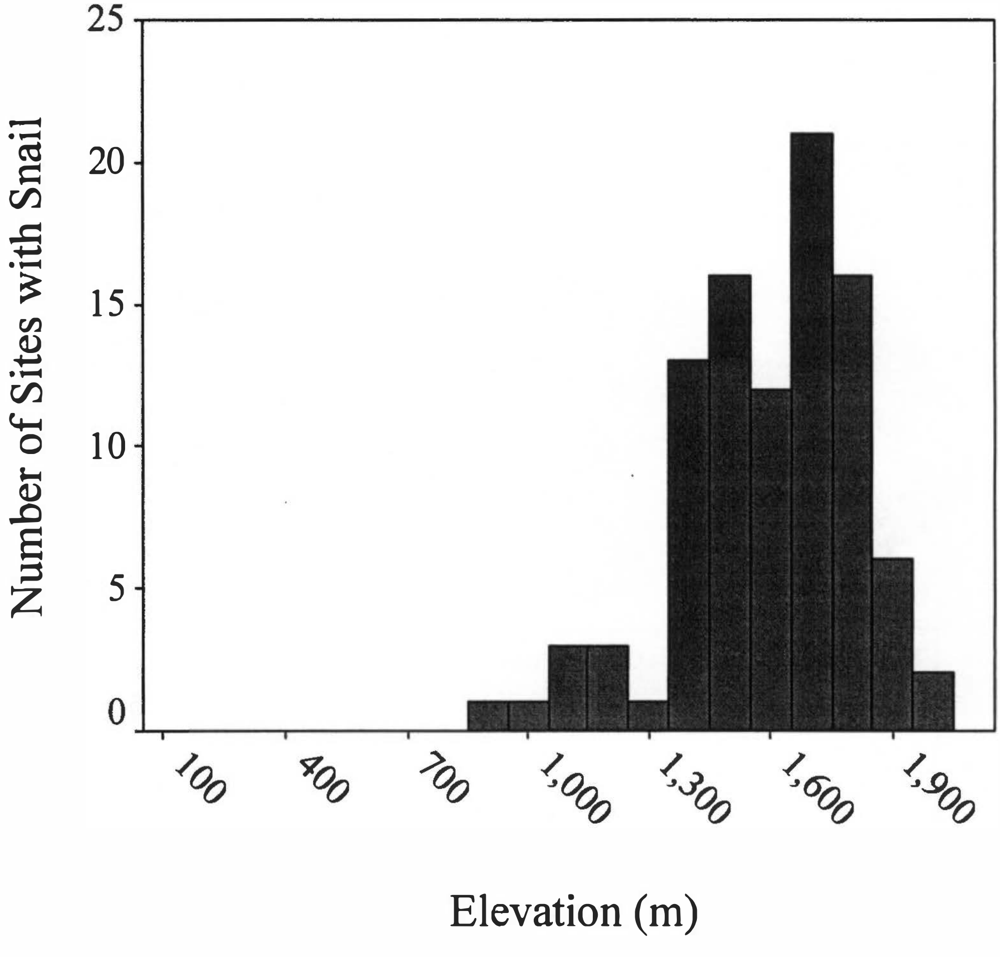


Fig. 5.4. Histogram of elevation (m) and presence of *V. latissimus.* Standard Deviation = 218.43, Mean = 1589, and n = 95.

5.4 Habitat Suitability Model Results

The final equation used to evaluate the area of suitable habitat (H) for *V. latissimus* in the park was:

$$H = (\text{elevation} * 0.50) + (\text{soils} * 0.30) + (\text{vegetation} * 0.20)$$

The combination of the reclassified environmental variables within ArcGIS resulted in a habitat suitability model (Fig. 5.5). The model shows the optimal habitat areas down to the least suitable habitat areas for *V. latissimus* in GSMNP. The final habitat map shows the extent of each ranked suitability, zones one through five. Class five, the most suitable habitat zone, covers 13.8% (8,873 ha) of the park, primarily in areas of high elevation. Class four, the next most suitable habitat zone for *V. latissimus* covers 23.3% (14,981 ha) of the park. Class three covers 34.6% (22,246 ha) of the park representing the largest ranked zone. Class two covers 21.5% (13,824 ha) of the park, primarily in lower elevations. Class one represents the least suitable habitat area in the park. The zone covers only 6.7% (4,308 ha) of the park, found entirely in the lower elevations.

5.5 Habitat Suitability Model Verification

Presence/absence field data were overlaid with the habitat suitability layer in the GIS to verify the results (Fig. 5.6). Sites that contain the land snail *V. latissimus* are found almost entirely in zones four and five as would be expected. Sites that do not contain the species are primarily in zones one, two, and three. The map overlay of the presence/absence data supports the validity of the model. Snails are found in the zones that have a high habitat suitability rank and the snail is not found where there are low

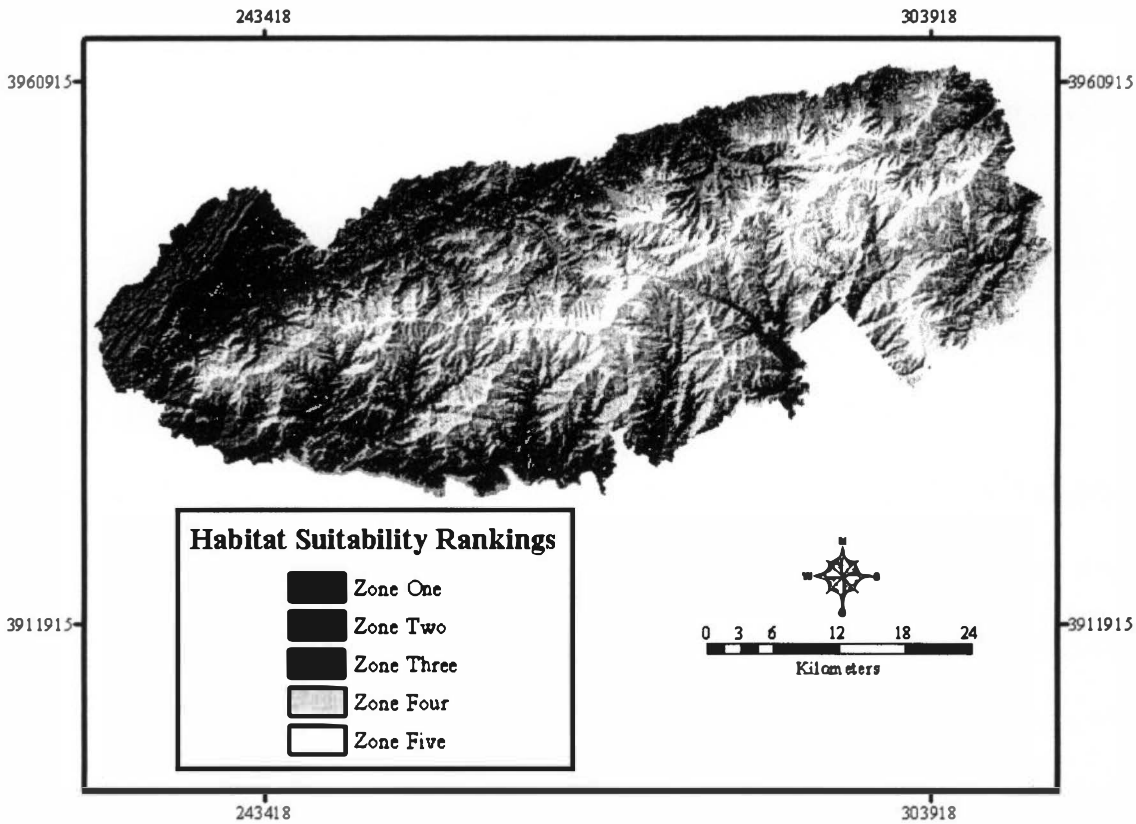


Fig. 5.5. Habitat suitability model for *V. latissimus.*

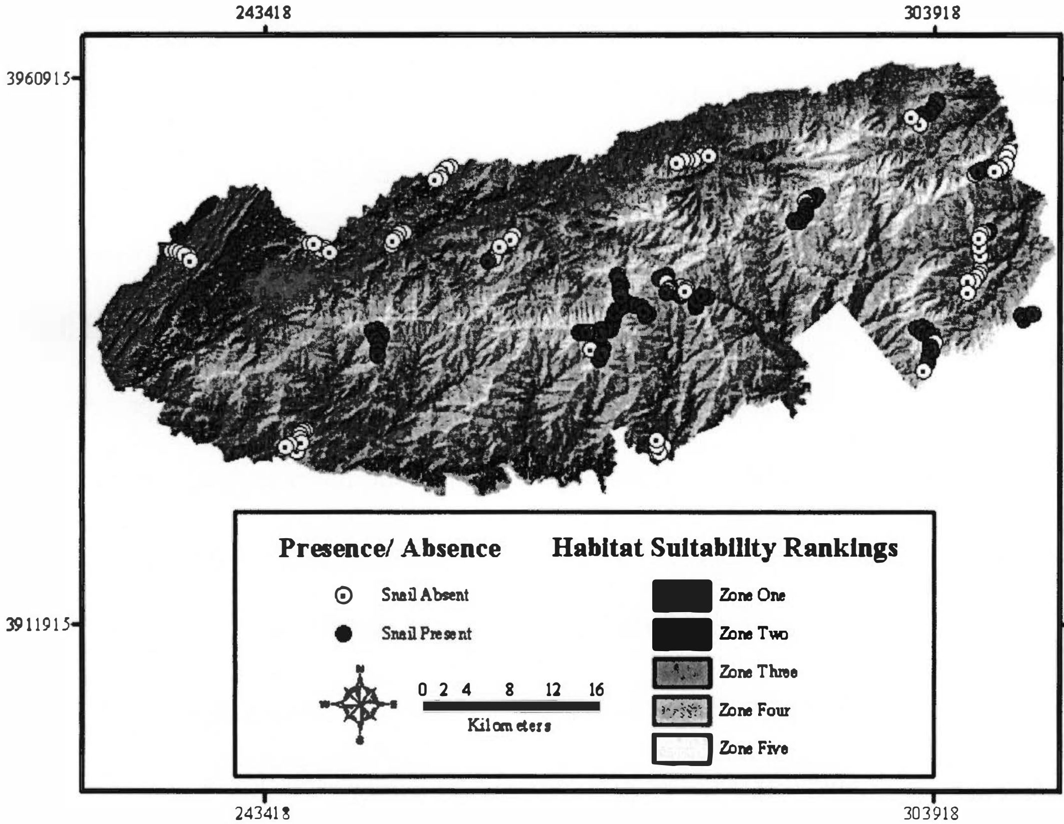


Fig 5.6. Verification map of suitability zones and the presence/absence of *V. latissimus*. Dark circles indicate the presence of the species. The positive occurrences of the species coincide with zones four and five of the habitat suitability model. Snails are not found in zones one and two and few occurrences happen in zone three.

habitat suitability zones. Further sampling for *V. latissimus* will likely show continued verification of the model.

5.6 Possible Extension of the Model

A preliminary example of extending the model for conservation investigation is offered in the form of species abundance models that compare *Vitrinizonites latissimus* and *Stenotrema depilatum* in the Clingman's Dome area of GSMNP. This area had the highest concentration of both species present in the study. *V. latissimus* has a protein-based shell and *S. depilatum*, like most land snails in the park, has a shell comprised of calcium. Less soil calcium is needed for *V. latissimus* to sustain healthy populations, but calcium leaching caused by acid deposition could hinder the health of *S. depilatum* populations in these areas. The abundance model shows a higher concentration of *V. latissimus* in the upper elevations of the park (Fig. 5.7).

No immediate relationships were observed between sites with more occurrences and significant environmental variables. The number of snails present did not vary by different soil types or vegetation. All sites were located in high elevations and varied from no *V. latissimus* to ten *V. latissimus* snails. Only five sites in the area had *S. depilatum* present, and each of these sites had two snails present. The preliminary results suggest that the protein-based *V. latissimus* snail is more abundant than the calcium-based *S. depilatum*. This may be the result of calcium leaching, but further research is needed to confirm this theory. Relationships between species richness and species abundance in acid deposition areas can also be analyzed when more data are collected on land snails.

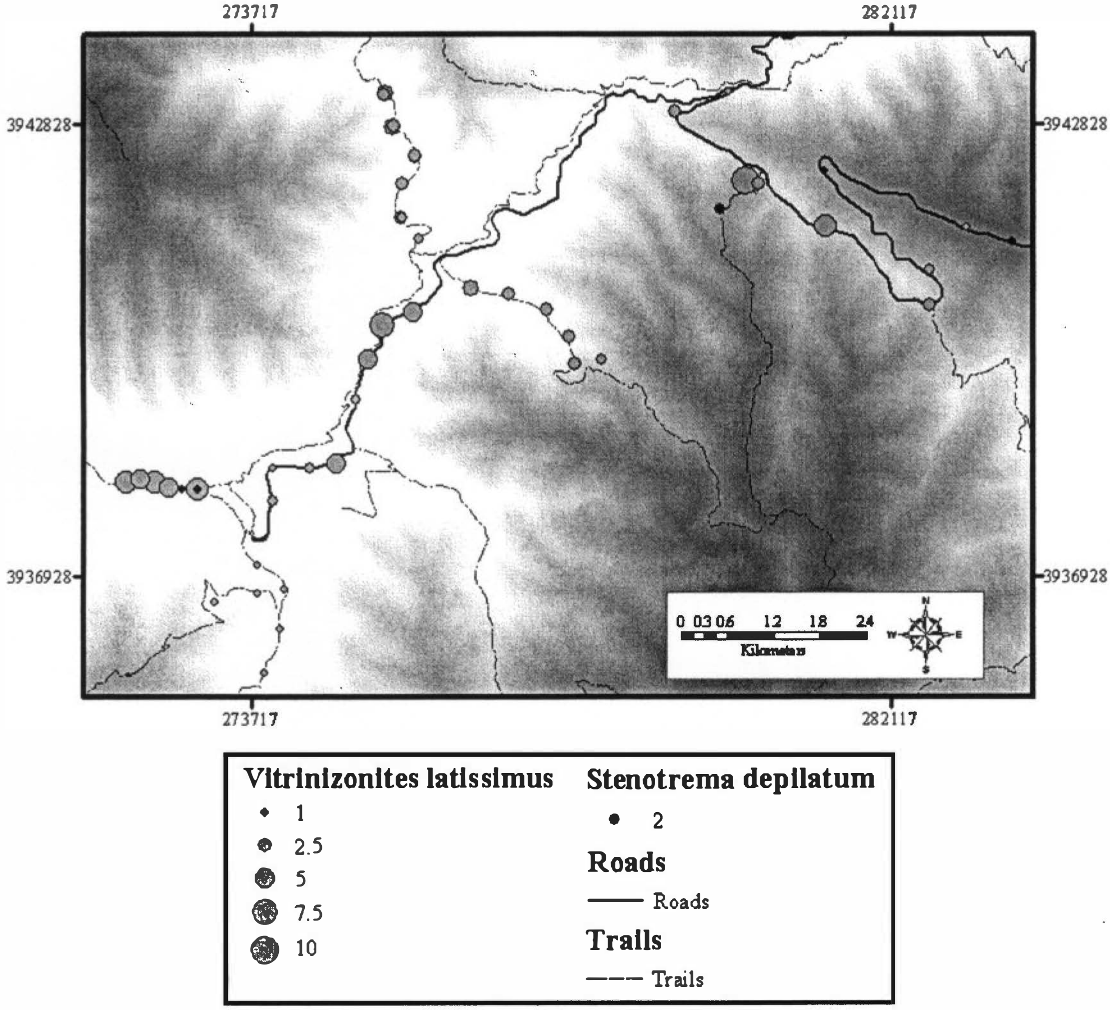


Fig. 5.7. Species abundance map for Clingman's Dome, GSMNP. Far more sites contain *V. latissimus* than *S. depilatum* in the Clingman's Dome area of the park. *V. latissimus* is also found in greater numbers.

Chapter Six
Discussion

6.1 The Habitat Suitability Model

Elevation is the primary factor that determines the distribution of *Vitrinizonites latissimus*, while soil and vegetation characteristics in its preferred high-country environments are secondary factors. The spruce-fir and northern hardwood forests are present in the majority of the most suitable habitat areas for *V. latissimus*. The humic matter in these vegetation areas upon which the snail feeds may contain more protein and preferred nutrients than humic matter from other vegetation types. *Rhododendron* spp. is typically less abundant in the spruce-fir and northern hardwood zones than it is in middle to low elevation areas of the park. Soils where *Rhododendron* spp. is present are typically more acidic and less nutrient-rich than other soil types. The plant absorbs nutrients (such as calcium) from the soil. Therefore, low and middle elevation areas that contain *Rhododendron* spp. would be a less suitable habitat for *V. latissimus*.

A study conducted in the high elevations of the Southern Appalachians shows a similar situation occurs with sugar maple, a species found in northern hardwood forests (Hotopp 2002). The study found that sugar maple was positively associated with land snail density and the amount of soil calcium present. The same study showed that red maple, a lower elevation tree type, has a negative association with land snail densities and available soil calcium. The study supports *V. latissimus* occurrences in northern hardwood forests and its absence in vegetation areas that include a high abundance of red maple. The study therefore suggests that elevation may indicate presence/absence of *V.*

latissimus as long as certain tree species are present. Further investigation into specific tree species and the occurrence of the snail could yield interesting results concerning the spatial distribution of these snails.

The acid deposition that occurs in the high elevations of the Smokies causes less calcium to be present in the soil and is detrimental to the abundance of most snails (Walden *et al.* 1991). Different species of snails have different requirements for calcium intake from their food. *V. latissimus* may thrive in areas of higher acid deposition rates because the species requires less calcium than other snails to survive. Soil acidification caused by acid rain has been shown to have a negative effect on snails and snail population densities. Snail species that depend on high calcium intake decline in areas of high acid deposition (Strom 2004). The decline of other snail species in these areas could cause *V. latissimus* to occur in greater numbers because competition for the available resources would be reduced. Further studies in GSMNP into the effects of acid deposition could support or refute these theories. By delineating the areas of acid deposition and studying snail samples from these areas, conclusions could be made about the effects acid deposition has on snail populations. Based on the results from this study, greater numbers of *V. latissimus* should be found in such areas.

Acid deposition has an adverse effect on both vegetation and soils, which are two significant variables for predicting the presence of *V. latissimus*. The snail may not be as affected by the leaching of calcium from the soil as are other snail species, but a change in soil chemistry and vegetation could still have an effect. Several studies were conducted on acid deposition in the high elevations of GSMNP by the Southern

Appalachian Mountains Initiative (SAMI 2002). They found that spruce-fir forests are more susceptible to the effects of acid deposition than other forest types. The tree becomes weaker and less tolerant to the cold winter conditions of the region. The needles of the spruce trees fall off prematurely which would affect the organic composition of the forest floor as well as cause more sunlight to reach the ground. The snail depends on the organic matter as a food source, and more sunlight would cause drier and less suitable conditions for *V. latissimus*. Degradation of the spruce-fir forest community could cause a decline in *V. latissimus* populations.

Increased levels of aluminum and mercury in the soils of GSMNP are caused by acid deposition and the metals are toxic to many tree and shrub species (Brewer *et al.* 2001). The metals inhibit the uptake of nutrients by plants and cause decreased root growth and root damage, causing a decline in biomass production. A loss in the abundance of the organic matter present would cause the food source for the snail to decline. If the organic matter becomes toxic, a decline in the snail population might result. Bioaccumulation becomes a problem as well. Organisms that feed on *V. latissimus* would have greater mercury intake that would continue through higher trophic levels, causing a decrease in the health of many organisms within GSMNP.

Most of the soils in GSMNP hold moisture relatively well. Soil moisture is a key factor in the survival of a snail, especially for *V. latissimus*, which does not have the ability to be enclosed completely within its shell to remain moist. The suitable habitat areas receive large amounts of rain and are wet the majority of the time. In areas of dense vegetation where sunlight is not as direct, the forest floor remains moist longer and is suitable for the moisture-dependent snails. The soils associated with *V. latissimus* are low in plant nutrients to begin with, and acid deposition causes them to be further deprived of nutrients. *V. latissimus* appears to be able to survive with a limited amount of plant nutrients which would explain the relative abundance of the species in high elevations of the park. A closer look at the feeding habits and nutrient intake of *V. latissimus* could help determine why these soil types are appropriate for the snail. It would also be useful to investigate which nutrients are most important to *V. latissimus.*

Land snails are well known for their poor mobility and are dependent on a stable microclimate that provides suitable habitat conditions. Temperatures are cooler and precipitation levels are greater in the higher elevations of the park. The greatest abundance of *V. latissimus* was observed on rainy or moist days or on days just before or after a rain event. Fewer snails were observed on hot, dry days. This would suggest that the snail is found in this region due to the greater frequency of cool, moist days in the upper elevations of the park. *V. latissimus* may require these climatic conditions to survive and this could be a reason for the influence elevation has on the species. The species may also thrive in lower elevation areas farther north that have similar temperature and precipitation levels as found in the high elevations of GSMNP. These hypotheses could be tested by conducting a similar study in low elevation areas in the northern extent of the species, where climate and vegetation conditions may be similar to the high elevations of GSMNP.

6.2 Habitat Variables Not Included in the Model

Based on the results of the cross tabulation and the subsequent logistic regression test of significance, the geology variable was not included in the model. Relationships were observed between the geology variable classes and the presence/absence data, but the logistic regression analysis suggested no significant relationship with geological substrate. The geology variable had one of the lowest logistic regression scores of any of the variables. The current geological data, however, are not as detailed as the data that are to be released in the near future. For example, the geology for a large area in the southeastern region of the park remains under study and has been temporarily classified as belonging to the Great Smokies Group. Geologists from the USGS are in the process of developing more complete and higher resolution geological data for the park to support ATBI efforts (Shultz *et al.* 1999). The newer, higher resolution geological substrate data may reveal that geology is in fact a significant variable for *V. latissimus.*

Only one logging history class was found to have high occurrences of *V. latissimus*, with this class being found primarily in the high elevations of the park. Based on the cross tabulation and logistic regression results, I found no relationship between the presence of *V. latissimus* and the logging class present at the site. Class five, the only class that had a positive residual value in the cross-tabulation results, represents areas of undisturbed forest. This class may have high counts of *V. latissimus* because of the average elevation in these areas. These areas are found primarily on steep slopes and high elevations of the park, and elevation is a significant factor. A study on land snail populations and forest regeneration in Sweden showed that logging history affected the distribution of certain land snail populations (Strom 2004). A similar investigation conducted in New England found no relationship between land snails and the logging history of that area (Strayer *et al.* 1986).

The continuous variables slope and aspect were also not included in the model. The t-tests conducted for these variables showed that they were non-significant for the presence of *V. latissimus*. Snails were observed on a variety of different slope values and a variety of aspects, and no associations were observed with the presence of *V. latissimus*. Some of the extreme slopes were not sampled, as they were deemed too hazardous. Slope plays an important role in soil moisture retention, which could possibly affect the distribution of land snails. I observed that fewer snails were present in drier areas that received more sunlight throughout the day. Aspect and vegetation coverage are the primary factors that regulate the amount of sunlight the forest floor receives.

Slope and aspect were not significant by themselves, but if they were combined in some way to represent soil moisture, they could be found to be significant indicators in future habitat studies. Soil moisture is a major factor in determining the distribution of land snails. If low amounts of soil moisture exist at a site, a snail will enclose itself in its shell and secrete a layer of mucus to remain moist (Baker 1939). *V. latissimus* has a body larger than its shell and can not enclose itself completely within its shell when there is not a suitable amount of soil moisture present. Therefore, soil moisture should be a determining factor for the presence of *V. latissimus*. Other organisms such as salamanders also require moist soil to survive. A GIS layer that contains data for soil moisture in the park would be a beneficial addition to the habitat model for *V. latissimus*, as well as habitat models for other snail species and moisture-dependent organisms.

6.3 Importance to Management

The habitat suitability model for *V. latissimus* helped determine the likely spatial distributions of the species throughout GSMNP. Habitat models of this nature can help researchers visualize vast datasets and provide a better understanding of the spatial distributions of species. Typically, distribution and habitat models consist of maps with simple dots that represent the known occurrences of a species. The model developed for this study provides environmental explanations concerning why these dots occur where they do. Investigators can see beyond the (x,y) location of the species occurrence and determine which environmental variables are present that may influence the distribution of the species. Once the habitat characteristics are determined, further investigations can be made concerning why these factors influence the distribution of the species.

As more biological data become available through the ATBI, such models will provide an improved understanding of the suitable habitats in the park for certain species. Researchers should be able to implement this habitat suitability model approach in their own research for a variety of organisms found in GSMNP. As more environmental variable data become available, more variables can be included to redefine the model. This model was designed with the ATBI research efforts in mind. This model utilizes GIS software without complex manipulations or complicated processes that would hinder a user unaccustomed to GIS from making use of the technology. ATBI scientists should benefit from the procedures outlined in this study.

One of the major goals of the ATBI is to produce distribution maps, which are the primary result of the habitat suitability model procedures outlined in this study. Based on the samples collected in the field, the model supplements the distribution information about a species. The model shows areas throughout the park that should contain the species. Similar models can be created for selected species to represent the geographic extent that the organism covers in the park. This aspect allows researchers that are studying organisms that have been previously modeled to focus their work in sections of the park that should contain their species in relative abundance.

Compiling life history data (*e.g.* feeding habits, reproduction, physiology, distribution), another goal of the ATBI, is supported by this habitat suitability model approach. Habitat characteristics are derived through the model and further life history investigations can be undertaken with the implementation of the model. Investigations into food sources, organism size, and other aspects describing the organism can be undertaken once the spatial distributions have been more clearly defined by the model. Life history research results can be managed through the geodatabase and site comparisons can show differences in life history characteristics such as difference in size or feeding habits between each site. For example, does the physical size of *V. latissimus* vary between occurrences in spruce-fir forests and northern hardwood forests? The database for the organism could be queried to efficiently investigate such questions. Occurrences in northern hardwood forest could be selected from the data table and exported as a new GIS layer. A new layer would also be created for spruce-fir forests. Graphs could be produced to show shell size by forest type and any differences between the vegetation present and the size of the shell could be observed.

A long-term goal of the ATBI is to determine a species' role in the greater ecosystem of GSMNP. Each organism in the park plays a role in the greater ecosystem and often has an effect on the distribution of organisms in higher trophic levels. The integration of several habitat suitability models for different species can provide a visual representation of how different components of the GSMNP ecosystem interact with each other. Distributions of two different organisms can be overlaid within the GIS and queried to determine if the organisms benefit from a shared resource. For example, if an overlay of rhododendron and eastern hemlock show that the two species occur together in the majority of the park, habitat variables (*e.g.* slope, aspect, elevation, soils, and geology) can be analyzed to determine the shared characteristics of these plant species and investigate why they are found in association with each other. Predator/prey relationships can also be utilized to enhance the habitat suitability models. For example, models of prey species can be used as a variable in the habitat suitability model of its predator. Areas that have a greater abundance of snails are more likely to contain black-capped chickadees, which feed on the snails. The suitable habitat layer for the snail species could be used as an environmental variable for the habitat suitability model of the black-capped chickadee. Relationships among the distributions of different organisms will become evident, thus helping the ATBI meet one of its long-term goals.

Implementation of GIS-based habitat models in ATBI research could enhance conservation efforts in the park. Habitat models may be integrated with each other to help determine which characteristics of the park are most important for focusing research and conservation development strategies. For example, the habitat of every modeled species can be combined within the GIS to determine the areas that contain the most diverse biology in the park, and the data can then be queried to determine the most common characteristics of these areas. The habitat layer for each species would be added together in the GIS raster calculator and the output would contain rankings similar to that of the habitat suitability model. The new layer would identify the areas of the park that contain the greatest biodiversity. Higher numbers would indicate a more diverse area. Park conservationists can focus their efforts on preserving diversity in GSMNP by querying the data to determine which environmental characteristics are present in these areas of high diversity. The resulting areas would be given higher conservation priority.

Knowing habitat characteristics of organisms in the park can help determine the effects of environmental degradation on the park ecosystem caused by invasive species. A GIS-based predictive model may be able to investigate which organisms will be adversely affected by the loss of a particular species of tree due to an invasive species. If the tree species is a determining environmental variable for an organism, the loss of the tree will affect the organism's distribution and abundance in the park. For example, Fraser fir, a species currently threatened by the balsam wooly adelgid, is a tree species included in the habitat characteristics of *V. latissimus* (Sharkey 2001). Pixel values in the vegetation layer representing this tree species can be removed from the layer and replaced with a treeless value for the pixel. The habitat suitability model can be reevaluated based on the new vegetation information. The output of the model would show the organism's distribution, reflecting the loss of the Fraser fir. The new model can be compared to the previous model to determine how the balsam wooly adelgid will affect the distribution of *V. latissimus.* This information can be important for making informed management decisions that concern the expected spread of an invasive species.

The use of this habitat suitability model approach will help park administrators assess the possible effects of increased acid deposition levels caused by air pollution on GSMNP organisms, and can provide park personnel with supplementary information as they fight for tighter air quality regulations. The habitat suitability model is the key factor in GIS-based predictive environmental response models (Goodchild *et al.* 1983). An organism's response to the loss of other key species in the park caused by air pollution can be modeled using various habitat suitability models and the results can be applied to help prevent further loss. For example, the red spruce is a species sensitive to air pollution and acid rain (NPS 2001). Like the Fraser fir, the red spruce is a tree included among the preferred habitat characteristics of *V. latissimus.* The loss or decline of the red spruce can be modeled to predict the effect it will have on the land snail.

Not all red spruce stands, however, are affected by air pollution. Park resource managers could delineate the extent of the park most affected by air pollution. The GIS could identify which stands of red spruce are likely to be affected by increased air pollution by spatially selecting within the GIS all red spruce stands that occur in the delineated air pollution zones. The pixel values for these specific stands could be changed to treeless and the model could be reevaluated based on this new information. The result of the model could then be compared to the normal habitat model to predict the effect increased air pollution will have on *V. latissimus*. This comparison could be used by park officials to support the demand for tighter air quality regulations. The ability to implement these predictive models allows conservation planners to better prepare for the future of GSMNP.

The habitat suitability and abundance model procedures will benefit logistic planning in the park. The impacts of new road construction or maintenance may become evident by assessing the abundance of certain species that are found within a proposed road corridor. A GIS-based model will help determine which path can impose the least biological impact by analyzing the overlap of habitat suitability models and the proposed road corridor. If the corridor contains optimal habitat zones for several species, the roadway may need to be rerouted to avoid such species-rich areas. For example, if a new road was to be constructed along the western shore of Fontana Lake, located on the southeast border of the park, a GIS could help determine the route that would impose the least biological impact. The proposed road corridor would be digitized as a GIS layer and given a buffer of 20 m on either side to represent the area impaired by the construction. Habitat models that had been generated for individual GSMNP organisms would be clipped to the extent of the road buffer using the geoprocessing command in ArcGIS. The individual models would be combined in the raster calculator to visualize the diversity found within the road corridor. High numerical values within the buffered zone would indicate that the area is a suitable habitat for a number of species and should be avoided. If the numerical values are low within the buffered zone, the new road would have little impact on the diversity of the area depending on which species are present. GIS layers that represent factors key to road construction (such as elevation, slope, and geology) can be reclassified to the degree they are suitable for a road. These reclassified layers and the new diversity layer could be evaluated by the raster calculator with weights given to the most important factors. The output would specify a linear feature representing the optimal path for the new road based on species diversity and available site conditions. In this manner, GIS models can support logistic planning in the park to avoid environmental degradation when developing roads or similar features.

The habitat models can also be a boost for tourism in GSMNP. The models may be implemented with GIS-based Spatial Decision Support Systems (SDSS) which are used to help visitors plan their excursions in the park. A SDSS offers the user a list of factors and constraints to choose and rank according to their personal interests for their trip. The habitat models can be entered as factors. If a visitor would like to encounter certain organisms during their hike, the system can prioritize trails that coincide with optimal habitat areas for that organism. For example, if the user would like to have an encounter with the diverse population of salamanders in the park, they would rank salamanders as a high priority in the SDSS. The system would output a list of trails that travel through the optimal habitat areas for salamanders and would give the user the best chance for encountering a salamander. If the visitor wants to see a variety of wildlife, they would rank several different species as a high priority and the SDSS would output trails that traveled through the areas with the greatest suitable habitat overlap.

Once the ATBI has been completed, all species found within GSMNP will be inventoried. The habitat models will be able to determine where select species in the park are found and why they are found where they are. If the ATBI is successful in GSMNP, the inventory methods can be implemented in National Parks around the nation (Sharkey 2001). The GIS-based habitat models may become a useful tool in conservation throughout the National Park system. This research offers a starting point to build upon, and as more data are gathered, sampling techniques are improved, and GIS technology advances, these models will offer a vital tool in the future of conservation management.

6.4 Statistical Considerations

Several previous studies used logistic regression to identify relationships between the presence of a species and environmental variables. Logistic regression could not be used as the primary statistical test in this study because significant colinearity exists among the environmental variables used (Fig. 6.1, 6.2, and 6.3). For example, elevation and geology are related. Therefore, logistic regression using multiple variables could not be used to identify relationships in this study with the presence/absence data. The technique was still useful, however, in determining weights for the model.

Previous studies outside GSMNP were able to use logistic regression to identify relationships between variables because they used datasets that did not contain significant colinearity. For example, a study of gray wolf habitat in the northern Rockies used land cover, elevation, aspect, and ownership type in a logistic regression model to show significance (Brehme 2001). These variables did not contain colinearity. Studies that

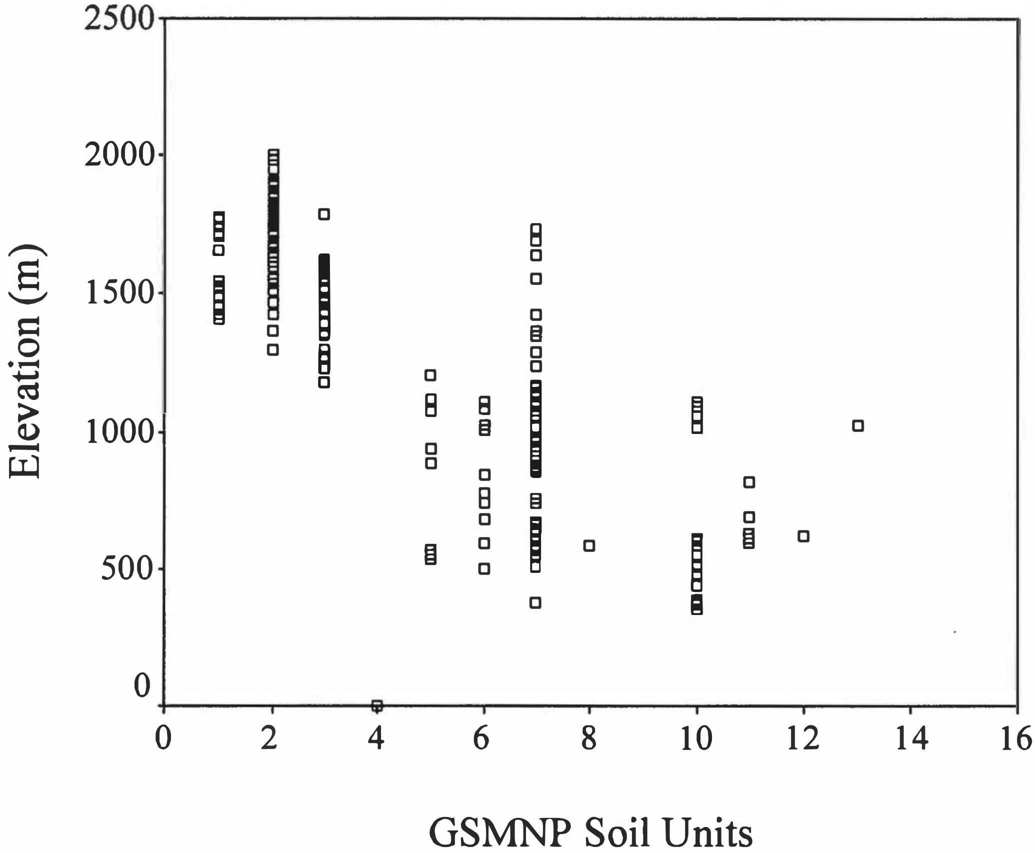


Fig. 6.1. Scatterplot of elevation (m) and soil units of GSMNP. The graph represents sites at which snails were found. See Table 3.1 for an explanation of the soil units.

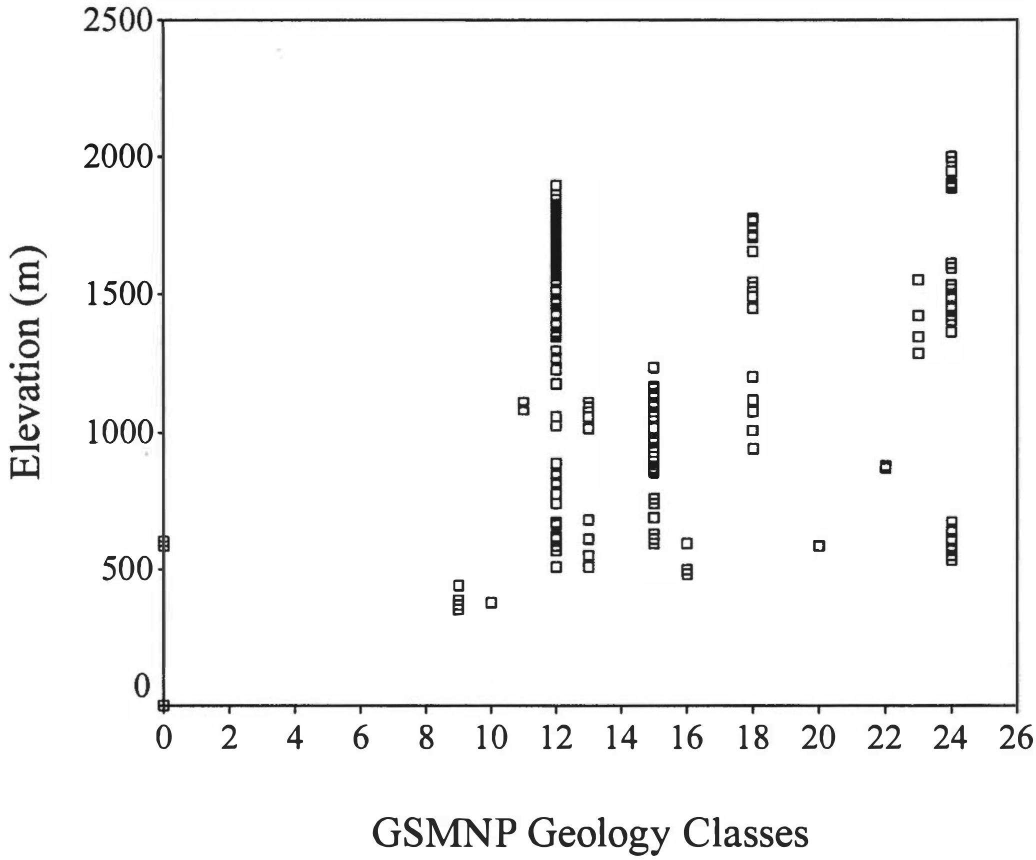


Fig. 6.2. Scatterplot of elevation (m) and geology classes in GSMNP. The graph represents sites at which snails were found. Geology class 1 is Blockhouse shale, class 2 is Lenoir limestone, class 3 is limestone/dolomite, class 4 is Cochoran formation, class 5 is Nichols shale, class 6 is Nebo quartzite, class 7 is Murray shale, class 8 is Hesse quartzite, class 9 is Wilhite formation, class 10 is Wilhite formation coarse, class 11 is Cades sandstone, class 12 is Thunderhead sandstone, class 13 is Metcalf phyllite, class 14 is Pigeon siltstone, class 15 is Roaring Fork sandstone, class 16 is Elkmont sandstone, class 17 is Wading Branch formation, class 18 is Anakeesta formation, class 19 is unnamed sandstone, class 20 is basement complex, class 21 is Metadiorite, class 22 is Longarm quartzite, class 23 is Rich Butt sandstone, class 24 is Great Smoky Group, and class 25 is Shields formation.

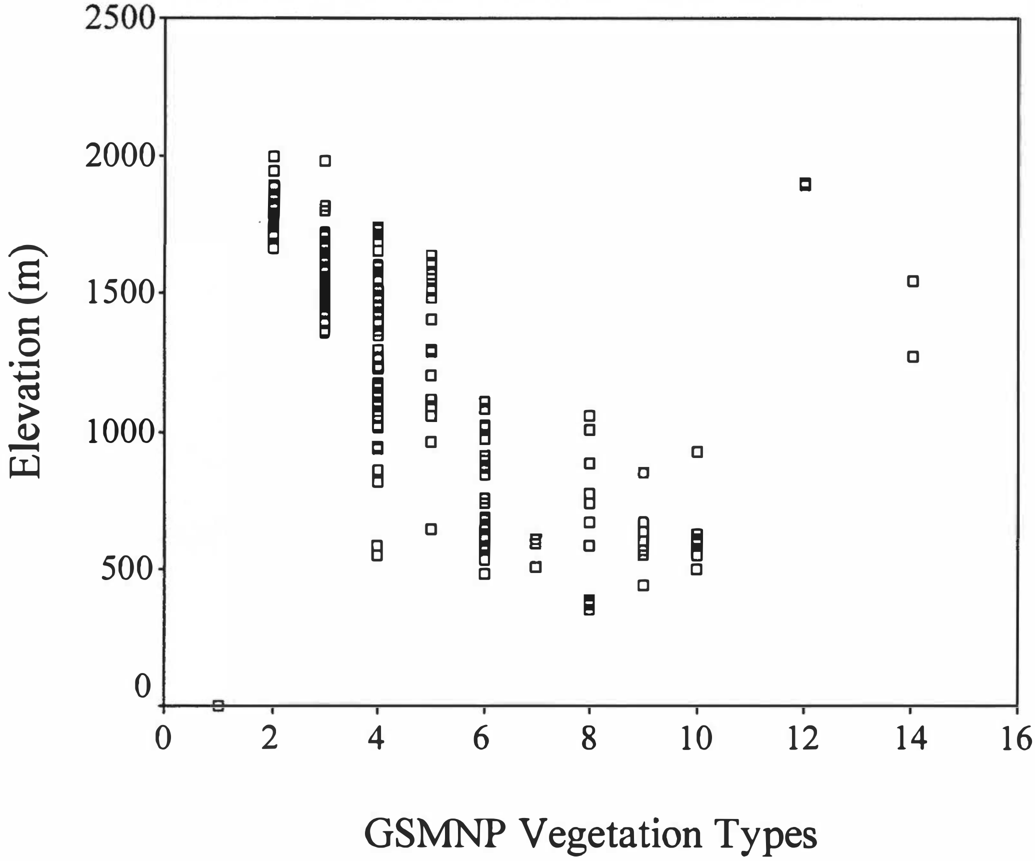


Fig. 6.3. Scatterplot of elevation (m) and vegetation types in GSMNP. The graph represents sites at which snails were found. Vegetation type 1 is spruce-fir, type 2 is northern hardwood, type 3 is cove hardwood, type 4 is mesic oak, type 5 is mixed mesic hardwood, type 6 is tulip poplar, type 7 is xeric oak, type 8 is pine-oak, type 9 is pine, type 10 is heath bald, type 11 is grassy bald, type 12 is grape thicket, type 13 is treeless, and type 14 is water.

analyze data that contain colinearity use different methods to test for significance. For example, a study that determined suitable habitat areas for prairie dogs in the Northern Great Plains used a classification tree as the primary test of significance to determine which variables were associated with the animal (Proctor *et al.* 1998). Some studies still use logistic regression for their analysis despite the colinearity present in the data. For example, a study of wood thrush habitat in GSMNP used elevation and geology among other variables in their model. Elevation and geology show significant colinearity (Fig. 6.2) and should not be used in the same logistic regression model (Shriner *et al.* 2002).

6.5 Possible Sources of Error

Several sources of error were possible in this study. Human error is often a major source of error in studies that involve fieldwork and manual data input. First, it is possible that some species may have been misidentified in the field during the first few sample runs while I was still learning how to identify the species. Second, the temporal variation in the weather over the course of sampling may have been a factor in the snail abundance observed. Samples taken in colder months or on drier days may have yielded fewer occurrences than samples on relatively warm or wet days. The samples were obtained over a span of eight months under a variety of weather conditions. Sites could be sampled multiple times during the course of the study and taken under different weather conditions to avoid the temporal variation between sample sites.

Other human errors may have occurred in the laboratory during analysis. When manually adding the continuous variables of slope, aspect, and elevation, some values may have been entered incorrectly. For example, if the sample site symbol contained more than one pixel, values may have been taken from the wrong pixel and entered incorrectly into the data table. This can be avoided by shrinking the symbol size for sample points and zooming to the sample point level to obtain the value of the pixel that contains the majority of the site.

The technology used in this study may have introduced additional sources of error. Although field measurements obtained with the GPS unit were corrected for error sources, measurement error may still have remained for some of the data that would affect accurate geographic placement. The GIS data may contain error as well. Polygon data contain distinct borders and do not represent the actual gradual transition from one attribute type to another. For example, an obvious border does not exist between two different vegetation types. Rather, a gradual transition zone occurs that contains elements of both types of vegetation along the border. The data layers are rough representations of reality. Error may occur in the original data obtained from outside sources and would join erroneous attribute data to the sample site to give false information for the geographic representation of the site. The resulting output is a function of the input values. Error in the data propagates to the output of the operation.

Nonetheless, such errors were likely small and would not have a great effect on the validity of the resulting output. Major errors that occurred were noticed and removed from the model before the operation was executed. Errors contained within the operation are likely insignificant and have little effect on the conclusions of this study.

Chapter Seven
Conclusions

7.1 Preferred Habitat Characteristics of *Vitrinizonites latissimus* in GSMNP

The habitat suitability model identified the preferred habitat characteristics for the land snail species *V. latissimus.* The following conclusions can be inferred about this snail's habitat:

1. Spruce-fir and northern hardwood forests are the preferred vegetation types of V. latissimus.

These forest types are found predominantly in the high elevations of the park and contain red spruce, Fraser fir, yellow birch, American beech, eastern hemlock, sugar maple, and black cherry. The humic matter in the soils in these vegetation types may contain more protein and preferred nutrients than soils in other vegetation types. Also, these areas contain less R*hododendron spp.* than other areas of the park. The rhododendron takes nutrients (such as calcium) from the soil, and may make the surrounding area less suitable for *V. latissimus* due to the change in soil chemistry.

Cove hardwood forests and grassy balds are secondary preferred habitats for *V. latissimus.* Trees in the cove hardwood forest types include sugar maple, a tree species observed to have a positive association with land snail densities and soil calcium (Hotopp 2002). The snail should be found with a relatively high probability, but soil type and soil moisture properties may also be important factors. The least suitable vegetation types were pine, xeric oak, tulip poplar, and pine oak forests, as well as mixed mesic hardwood forests and treeless areas. These tree types are found primarily in lower elevations of the park and in areas of non-suitable soil units for *V. latissimus*. These forest types also include red maple, a tree species observed to have a negative association with land snail densities and soil calcium (Hotopp 2002).

2. Breakneck-Pullback, Oconaluftee-Guyot Chiltoskie, and Luftee-Anakeesta are the most suitable and preferred soil units for V. latissimus.

The habitat of this species is most likely to consist of these soil units, and the snail should occur in these areas with a high frequency depending on the vegetation and elevation present within these soil units. Complicating this finding, however, are changes in soil chemistry due to acid deposition in these units that may have a greater effect on other snail species than on *V. latissimus*. The decline of these other snail species could reduce the competition for available resources and cause *V. latissimus* to occur in greater numbers. Most GSMNP soil units hold moisture relatively well. Furthermore, the greater amounts of precipitation in the higher elevations of the park cause soils to be moist for a greater proportion of time. Because *V. latissimus* is a moisture-dependent organism, it would be found in greater numbers in moister areas. Finally, the preferred soil types are also low in plant nutrients. *V. latissimus* appears to be able to survive with a limited amount of plant nutrients, which would explain their relative abundance in these soil units.

Secondary soil units for the habitat of *V. latissimus* include the Soco-Stecoah-Spivey, Dellwood-Smokemont-Reddies, and Wayah-Tanasee units. If the right combination of vegetation and elevation habitat components occur in these areas, *V. latissimus* may be found in these soil types as well. The least suitable soil classes for *V. latissimus* are the Junaluska-Tsali, Ditney-Unicoi-Spivey, Spivey-Santeetlah, Lauda-Fanning, and Evard-Cowee-Leatherwood units. These soil units are found in lower elevations of the park and most often do not contain the preferred vegetation types for *V. latissimus.*

3. The most suitable elevation range for the species is elevations above 1400 m.

Elevations above 1400 m are the principal habitat for *V. latissimus.* The preferred combination of soil units and vegetation types occurs most frequently in the higher elevations of GSMNP. Greater amounts of precipitation occur in the higher elevations of the park, which would cause more moisture to be present in the soils. Temperatures are cooler in these elevations and more snails were observed on relatively cool and moist days. *V. latissimus* may require the climatic conditions that occur in the higher elevations of the park to survive and this could be an additional reason that upper elevations of the park are a preferred habitat of the species.

7.2 Distribution of *V. latissimus* in GSMNP

The habitat suitability model provided information on the land snail's probable spatial distribution in GSMNP. The following conclusions can be inferred:

1. V. latissimus is more likely to be present in zones four and five of the habitat suitability model.

Zone five consists of spruce-fir and northern hardwood forests, Breakneck-Pullback, Oconaluftee-Guyot Chiltoskie, and Luftee-Anakeesta soil units, and elevations above 1400 m. This zone represents 8,872 ha of optimal habitat conditions for *V. latissimus* in GSMNP. Zone four also consists of spruce-fir and northern hardwood forests and also includes cove hardwood forests as a vegetation component. In addition to soil units also found in zone five, zone four also contains Soco-Stecoah-Spivey, Dellwood-Smokemont-Reddies, and Wayah-Tanasee soil units. The 14,981 ha of the park contained in zone four are found at elevations ranging between 1100 m and 1600 m. These higher elevations are found primarily on the center ridgeline of the park along the border of Tennessee and North Carolina, especially at Clingman's Dome and in the northeast portion of the park around Balsam Mountain. Optimal zones also spur off along other high ridges of GSMNP along its southern border.

2. The species is less likely to occur in zone three of the habitat suitability model.

Zone three consists of mesic oak and cove hardwood forests, Soco-Stecoah-Spivey, Dellwood-Smokemont-Reddies, and Wayah-Tanasee soil units located in mid to low elevation areas. Zone three contains 22,246 ha of the park, including large areas near Rich Mountain and Little Roundtop between Townsend, Tennesee and Gatlinburg, Tennessee.

3. Zones one and two of the model should contain no occurrences of V. latissimus.

Zones one and two consist of less suitable vegetation types which include pine, xeric oak, tulip poplar, pine-oak, mixed mesic hardwood, and treeless forest types. Less suitable soil types include the Junaluska-Tsali, Ditney-Unicoi-Spivey, Spivey-Santeetlah, Lauda-Fanning, and Evard-Cowee-Leatherwood units. These less desirable areas are found in 18,131 ha of the park. These zones are found in lower elevations of the park in the valleys, such as Cades Cove in the southwest and Cataloochee in the northeast. They are also found along the boundaries of the park in low-lying areas such as Fontana in the southeast, Little Bottoms in the southwest, and along the Little River that runs from Gatlinburg, Tennessee to Townsend, Tennessee.

7.3 The GIS-based Model

1. GIS technology can be implemented as a useful tool in modeling probable habitat distributions for organisms of GSMNP.

The habitat suitability model for *V. latissimus* provides an example of how GIS-based habitat models can be useful tools for biological research in GSMNP. Field data collected for the organism in GSMNP enabled the GIS to calculate the probable habitats for the snail by querying environmental data layers and combining the weighted, significant variables. Such models are useful in visualizing the probable distributions of organisms in GSMNP. This aspect helps researchers focus their study on areas that are most likely to contain their species. The model approach is also helpful for evaluating life history attributes of a species related to specific environmental variables. For example, characteristics of humic matter of the soil in one vegetation type might contain the preferred food source for a species while other vegetation types may have less suitable food sources available.

The model is an accurate depiction of the probable distribution of *V. latissimus* in GSMNP based on the current environmental conditions in the park. Further sampling should continue to prove the validity of the model. GIS technology is dynamic in nature. As more data are gathered, or as environmental conditions of the park change over time, the model should be able to incorporate the additional data or account for the new environmental conditions. For example, if a major decline of the spruce-fir forest type occurs, the data can be updated and the model can be rerun to represent the new environmental conditions.

2. The model is useful for conservation and management efforts within GSMNP.

The GIS-based technology will become increasingly helpful in supporting conservation and management issues for GSMNP. Combining habitat models for a variety of park organisms can produce a model of the most diverse areas of the park. Habitat features can be studied to determine why increased biodiversity exists in these areas. Furthermore, the effects of acid deposition and invasive species can also be added to the habitat suitability model. Tree species within acid deposition areas or under attack by an invasive species can be removed from the vegetation variable and the habitat model can be reevaluated to study how the new model reflects the loss.

Effects of developments (such as roads and parking lots) can be modeled to visualize how they will affect certain organisms and the diversity of the area. For example, a proposed roadway can be buffered by the GIS geoprocessing procedures and overlaid with the diversity model. Areas within the buffer that contain several areas of diversity should be avoided. Combining this information with variables that determine roadway suitability (geology, slope, and elevation) could provide a path that causes the least amount of biological impact.

Tourism within the park also can benefit from these GIS models. Visitors to the park can rank habitat models of popular species such as black bear (*Ursus americanus* Pallas), elk, and salamanders as factors in a GSMNP GIS-based SDSS. The system would offer the user a list of trails based on the user-specified factors and constraints, including the habitat model results, that best meet their personal interests. The system would offer the visitor an excursion that is unique to their requests.

3. The geodatabase environment is an efficient and viable option to store and manage GSMNP biological data.

The ArcGIS geodatabase environment proved to be an effective means for managing the data obtained during this study. Data were easily updated and organized within the geodatabase. Joining attributes from other data sources was efficiently executed within the geodatabase, and exporting the data to the SPSS software for statistical analysis was relatively simple. The dynamic nature of the biological data of GSMNP is easily maintained by the geodatabase. New datafields can easily be added and current biological datasets can easily be updated within the ArcCatalog geodatabase framework. Data that are collected in the future can be imported into the geodatabase and automatically given the correct coordinate system information. The geodatabase is able to handle the vast amounts of data that are to be collected by the ATBI.

Literature References

Adam, N.R., and A. Gangopadhyay. 1997. *Database Issues in Geographic Information Systems.* Kluwer Academic Publishers: Boston, Massachusetts.

Akçakaya, H.R. 1994. RAMAS/GIS: *Linking Landscape Data With Population Viability Analysis* (version 1.0). Applied Biomathematics: Setauket, New York.

Akçakaya, H.R. 1999. *Linking GIS with Models of Ecological Risk Assessment for Endangered Species.* Applied Biomathematics: Setauket, New York.

Baily, L.L. 2002. Estimating detection probabilities for terrestrial salamanders in Great Smoky Mountains National Park. North Carolina State University. Manuscript on file at NCSU library.

Baker, F.C. 1939. *Fieldbook of Illinois Land Snails.* State of Illinois: Urbana, Illinois.

Beltz, M., W. Haskins, and J. Proctor. 1998. A GIS Model for Identifying Potential Black-Tailed Prairie Dog Habitat in the Northern Great Plains Shortgrass Prairie. 1998 *ESRI International User's Conference Proceedings.* [Electronic Version]. Retrieved November 20, 2003 from http://gis.esri.com/library/userconf/proc98.

Bonham-Carter, G.F. 1994. *Geographic Information Systems for Geoscientists: Modelling with GIS.* Pergamon: New York, New York.

Binkley, D. 1989. *Acid Deposition and Forest Soils.* Springer-Verlag: New York, New York.

Blake, B., and S. Brown. 2001. Mapping Urban Deer Populations Using GPS and GIS. *2001 ESRI International User's Conference Proceedings.* [Electronic Version]. Retrieved November 20, 2003 from http://gis.esri.com/library/userconf/proc01.

Brady, W.W., and G.L. Whysong. 1999. Modeling. In Morain, S. *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

Brehme, C. 2001. A GIS Analysis and Model of Suitable Gray Wolf Habitat in the Northern Rockies. [Electronic Version]. Retrieved October 14, 2003 from http://www.spatial.maine.edu/ucgis.

Brewer, F.B., T.J. Sullivan, J. Cosby, R. Munson. 2001. Acid Deposition Effects to Forest and Streams in the Southern Appalachian Mountains. Southern Appalachian Initiative Report. [Electronic Version]. Retrieved March 3, 2004 from http://www.saminet.org/reports.

Burch, J.B. 1962. *How to Know the Eastern Land Snails.* Wm. C. Brown Company Publishers: Dubuque, Iowa.

Burt, J.E., and G.M. Barber. 1996. *Elementary Statistics for Geographers.* The Guilford Press: New York, New York.

Chang, Kang-tsung. 2002. *Introduction to Geographic Information Systems.* McGraw-Hill: New York, New York.

Chen, Yong-qi, and Yuk-Cheung Lee. 2001. *Geographic Data Acquisition.* Springer-Verlag: New York, New York.

Clarke, K.C., B.O. Parks, and M.P. Crane. (Eds.) 2002. *Geographic Information Systems and Environmental Modeling.* Prentice Hall, Inc.: New Jersey.

Couclelis, H. 2002. Modeling Frameworks, Paradigms and Approaches. In Clarke, K.C., B.O. Parks, and M.P. Crane. (Eds.) 2002. *Geographic Information Systems and Environmental Modeling.* Prentice Hall, Inc.: New Jersey

Craighead, L., and R. Walker. 1997. Least-Cost-Path Corridor Analysis Analyzing Wildlife Movement Corridors in Montana Using GIS. *1997 ESRI International User's Conference Proceedings.* [Electronic Version]. Retrieved November 20, 2003 from http://gis.esri.com/library/userconf/proc97.

Dalton, R., J. Garlick, R. Minshull, and A. Robinson. 1972. *Correlation Techniques in Geography.* George Philip and Son Limited: London, England.

Discover Life in America. 2003. http://www.dlia.org.

Dourson, D. 2002 to Andrew Dye, Graduate Student, University of Tennessee, Department of Geography. (September 28, 2002, 3:25 PM CST) personal correspondence, on file with recipient.

Egbert, S., A.T. Peterson, and K. Price. 1999. Modeling Conservation Priorities in Veracruz, Mexico. In Morain, S. *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

Falconer, A., and J. Foresman. 2002. *A System for Survival: GIS and Sustainable Development.* Environmental Systems Research Institute Press: Redlands, California.

Fedra, K. 1993. GIS and Environmental Modeling. In Goodchild, M.F., B.O. Parks, and L.T. Steyaert. (Eds.). *Environmental Modeling with GIS.* Oxford University Press: Oxford, England.

Galimberti, F., and S. Sanvito. 1999. GPS Mapping aids understanding of elephant seal behavior. *GPS World* 10: 22–30.

Gongbing, Peng, Lance M. Leslie, and Yaping Shao. (Eds.) 2002. *Environmental Modelling and Prediction.* Springer-Verlag: New York, New York.

Goodchild, M.F., B.O. Parks, and L.T. Steyaert. (Eds.) 1993. *Environmental Modeling with GIS.* Oxford University Press: Oxford, England.

Graveland, J.R. 1996. Avian eggshell formation in calcium-rich and calcium-poor habitats: importance of snail shells and anthropogenic calcium sources. *Canadian Journal of Zoology* 74: 1035–1044.

Graveland, J.R., R. vander Wal, J.H. van Balen, and A.J. van Noordwijk. 1994. Poor reproduction in forest passerines from decline of snail abundance on acidified soils. *Nature* 368: 446–448.

Great Smoky Mountains National Park. 2001. Orientation for Researchers in Great Smoky Mountains National Park. National Park Service. On file at Twin Creeks GSMNP.

Green, K. 1999. Development of the Spatial Domain in Resource Management. In Morain, S (ed.). *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

Groot, R., and J. McLaughlin. 2000. *Geospatial Data Infrastructure: Concepts, Cases, and Good Practice.* Oxford University Press: Oxford, England.

Gustafson, E.J., N.L. Murphy, and T.R. Crown. 2001. Using a GIS model to assess terrestrial salamander response to alternative forest management plans. *Journal of Environmental Management* 63: 281–292.

Guptill, S.C., and J.L. Morrison. 1995. *Elements of Spatial Data Quality.* Elsevier Science: New York, New York.

Hames, R.S., K.V. Rosenberg, J.D. Lowe, S.E. Barker, and A.A. Dhondt. 2002. Adverse effects of acid rain on the distribution of wood thrush *Hylocichla mustelina* in North America. *Proceedings of the National Academy of Sciences of the United States of America* 99: 11235–11240.

Havas, M. and T.C. Hutchinson. 1978. *Effects of Acid Precipitation on Terrestrial Ecosystems.* Plenum Press: New York, New York.

Heuvelink, G.B. 1998. *Error Propagation in Environmental Modelling with GIS.* Taylor and Francis: London, England.

Hofmann-Wellenhog, B., H. Lichtenegger, and J. Collins. 2001. *GPS Theory and Practice.* Springer-Verlag: New York, New York.

Hotopp, K. 2002. Land snails and soil calcium in central Appalachian mountain forests. *Southeastern Naturalist* 1(1): 27–44.

Johnson, C.A. 1993. Introduction to Quantitative Methods and Modeling in Community, Population, and Landscape Ecology. In Goodchild, M.F., B.O. Parks, and L.T. Steyaert. (Eds.). *Environmental Modeling with GIS.* Oxford University Press: Oxford, England.

Kemp, S. 1993. Trees and Familiar Shrubs of the Smokies. Great Smoky Mountains Natural History Association: Gatlinburg, Tennessee.

Kiser, A., and J. Williams. 2001. Mapping southern Oregon's noxious weeds with GPS. *GPS World* 11: 26–32.

Klopatek, J.M., and J.M. Francis. 1999. Spatial Pattern Analysis Techniques. In Morain, S. (Ed.). *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

Landgon, K. 2001. Personal Interview, September, Twin Creeks Gatlinburg, Tennessee.

MacDonald, A. 2001. *Building a Geodatabase.* Environmental Systems Research Institute Press: Redlands, California.

Morain, S. (Ed)1999. *GIS Solutions in Natural Resource Management: Balencing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

National Park Service. 2001. Nature and Science: Air Quality. U.S. Department of the Interior. [Electronic Version]. Retrieved April 12, 2002 from http://www.nps.gov/grsm.

National Park Service. 2001. Nature and Science: Non-Native Species. U.S. Department of the Interior. [Electronic Version]. Retrieved February 18. 2004 from http://www.nps.gov/grsm.

Natural Resources Conservation Services (NRCS) Soil Survey Staff. 1999. *Soil Taxonomy: A basic system of soil classification for making and interpreting soil surveys.* Second Edition. United States Department of Agriculture.

O'Connor, R.J., Paul R.H. Neville, and T.B. Bennett. 1999. Indicators of Resources and Landscapes. In Morain, S. (ed.) 1999. *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* OnWord Press: Santa Fe, New Mexico.

Peterson, A.T. 2001. Predicting species' geographic distributions based on ecological niche modeling. *The Condor* 103: 599–605.

Peterson, A.T. and A.G. Navarro-Sigüenza. 1999. Alternate species concepts as bases for determining priority conservation areas. *Conservation Biology* 13:427–431.

Peterson, A.T. , S.L. Egbert, V. Sanchez-Cordero, and K.P. Price. 2000. Geographic analysis of conservation priorities using distributional modeling and complementarity: endemic birds and mammals in Veracruz, Mexico. *Biological Conservation* 93: 85–94.

Pierce, D.S. 2000. The Great Smokies: From Natural Habitat to National Park. The University of Tennessee Press: Knoxville, Tennessee.

Peng, G., L.M. Leslie, and Y. Shao. 2002. *Environmental Modeling and Prediction.* Springer: New York, New York.

Price, M. and J. Tinant. 2001. Environmental parameters correlating to infestations of Canada Thistle in Badlands National Park: A GIS model. George Wright Society Biennial Conference, Denver, CO, April 2001.

Renfro, J. 2001 National Park Service Great Smoky Mountains National Park Briefing Statement. Manuscript on file at GSMNP.

Rogerson, P.A. 2001. *Statistical Methods for Geography.* SAGE Publications: London, England.

Sharkey, M. J. 2001. The All Taxa Biological Inventory of the Great Smoky Mountains National Park. *Florida Entomologist* 84(4): 556–564.

Shriner, S.A. 2001. Distribution of breeding birds in Great Smoky Mountains National Park. North Carolina State University. Manuscript on file at NCSU library.

Shriner, S.A, T.R. Simons, and G.L. Farnsworth. 2002. A GIS-based habitat model for Wood Thrush in Great Smoky Mountain National Park. *In Predicting Species Occurrences: Issues of Scale and Accuracy*. Scott, J.M., P.J. Heglund, M. Morrison, M. Raphael, J. Haufler, B. Wall, (Eds.) Island Press. Covello, California.

Shultz, A., and S. Southworth. 1999. USGS Open-File Report 99-175. *Digital Geologic Map of Parts of the Cades Cove and Calderwood Quadrangle, Tennessee and North Carolina, Great Smoky Mountains National Park.* [Electronic Version]. Retrieved September 20, 2003 from: http://geology.er.usgs.gov/eespteam/smoky/cades_cove/.

Simons, T. 2001. Assessing the Relationship Between Acid Precipitation, Calcium Depletion, and Avian Productivity in the Southern Appalachians. Pre-proposal Document. Manuscript on file at NCSU.

Slapcinsky, J. Florida Museum of Natural History, University of Florida, to Andrew Dye, Graduate Student, University of Tennessee, Department of Geography. (October 3, 2002, 5:47 PM CST) personal correspondence, on file with recipient.

Skidmore, A. 2002. *Environmental Modelling with GIS and Remote Sensing.* Taylor and Francis: London, England.

Solem, A., and A.C. van Bruggen. 1984. *World-Wide Snails.* E.J. Brill: Leiden, The Netherlands.

Southern Appalachian Mountains Initiative (SAMI). 2002. Final Report August 2002. [Electronic Version]. Retrieved February 7, 2004 from http://www.saminet.org.

Southworth, S. 2000. USGS Open-File Report 00-261. *Geology of the Mount LeConte 7.5-Minute Quadrangle, Great Smoky Mountains National Park, Tennessee and North Carolina.* [Electronic Version]. Retrieved September 20, 2003 from http://geology.er.usgs.gov/eespteam/Mtleconte/.

Stephenson, S.L., C.J. McQuattie, P.J. Edwards, and S.M. Studlar. 1992. The effect of acidification on the bryophytes associated with mountain streams. Page 66 in E. R. Smith (Ed.), Proceedings of the Third Annual Southern Appalachian Man and the Biosphere Conference, Tennessee Valley Authority, Norris, Tennessee.

Strayer, D., D.H. Pletsher, S.P. Hamburg, and S.C. Nordvin. 1986. The effects of forest disturbance on land gastropod communities in northern New England. *Canadian Journal of Zoology* 64: 2094–2098.

Strom, L. 2004. *Long-term Effects of Riparian Clear-cutting on Land Snail Communities in Regenerating Forests.* Master's Thesis. Department of Studies in Biology and Environmental Science. Umea University, Sweden.

Stupka, A. 1960. *Great Smoky Mountains National Park.* Natural History Handbook Series No. 5. United States Department of the Interior: Washington, D.C.

University of Idaho. 2002. Soil Science Division. http://soils.ag.uidaho.edu/soilorders/12orders.htm.

Varma, A. 2002. GIS Data Sources and Measurement Technologies for Modeling. In Clarke, K.C., B.O. Parks, and M.P. Crane. (Eds.) *Geographic Information Systems and Environmental Modeling.* Prentice Hall, Inc.: New Jersey.

Walden, H.W., U. Gardenfors, and L. Wareborn.1989. The impact of acid rain and heavy metals on the terrestrial mollusc fauna. Tenth International Malacological Congress, Tubingen: 425–435.

White, P.1996. *Wildflowers of the Smokies.* Great Smoky Mountains Natural History Association: Gatlinburg, Tennessee.

White, P. 2003. Speaker: 8th Annual ATBI Conference. December 8th Gatlinburg, Tennessee.

Wowrer, T.H. and R.G. Congalton. 2000. *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing.* Ann Arbor Press: Chelsea, Michigan.

Wuerthner, G. 2003. *Great Smoky Mountains.* Stackpole Books: Mechanicsburg, Pennsylvania.

Yankee, D. 1995. *Predictive Modeling of Deciduous Forest Stand Composition and Distribution Using GIS and Ecological Land Classification: a Regional-Scale Test in the Clinch-Powell Watershed of Tennessee.* Masters Thesis, The University of Tennessee, Department of Geography.

Zeiler, M. 1999. *Modelling Our World: The ESRI Guide to Geodatabase Design.* Environmental Systems Research Institute Press: Redlands, California.

Appendices

Appendix A
Vitrinizonites latissimus Description

Vitrinizonites latissimus Description

A.1 Taxonomy

- **Kingdom:** Animalia
- **Phylum:** Mollusca
- **Class:** Gastropda
- **Order:** Stylommatophora
- **Family:** Zonitidae

Mollusks comprise a large group of animals found in nearly all regions and all habitats throughout the world. Mollusks typically have a hard external shell; a mantle; which secretes the shell; soft bodies with no segmentation; mucus glands; and a large foot that may be variously modified for crawling or swimming. Six principle groups or classes of mollusks exist, most of which have names with reference to the foot, *i.e.*, gastropoda, or stomach-foot (Burch 1962).

The gastropods are the largest and most varied group of mollusks. They commonly have a flat ventral foot that is adapted for crawling, a coiled shell, and a coiled visceral mass. Unlike some other classes of mollusks, the gastropod's digestive tract is not a straight tube and the anus comes to lie in the side of the animal, often near the head. The largest order of gastropods is Stylommatophora. The animals of this group have two pairs of retractile tentacles with eyes at the tops of the upper pair (Burch 1962).

The Zonitidae family consists of small to medium snails and is found world-wide. The shell is usually perforate or umbilicate, and it typically has a depressed spire. The lip is thin and is not reflected. The margin of the foot is defined by a pedal groove, similar to other anulacopod snails. The body and foot of the snail are typically about twice as long as the diameter of the shell. The eye peduncles are lean and slender and the tentacles are short (Burch 1962).

A.2 General Description of the Species

Vitrinizonites latissimus is sometimes referred to by the common name glassy grapeskin, Blue Ridge snail, or the biting snail. The snail is found in the Southeast typically in the high elevations of Alabama, North Carolina, and Tennessee. The width of the snail is typically between 16.2–19.5 mm. The glossy shell is brown to olive, yellow, or green and has two and a half to three whorls. The spire is small, flat, and slightly sunken (Fig. A.1). The upper shell surface is marked with conspicuous, unevenly spaced radial grooves, which fade out at the shell periphery (Burch 1962). The shell is not large enough to completely contain the animal within (Fig. A.2). The shell is constructed of protein instead of calcium (Dourson 2002). The snail feeds primarily on leaf litter, bark, and other vegetative and humic material on the forest floor.

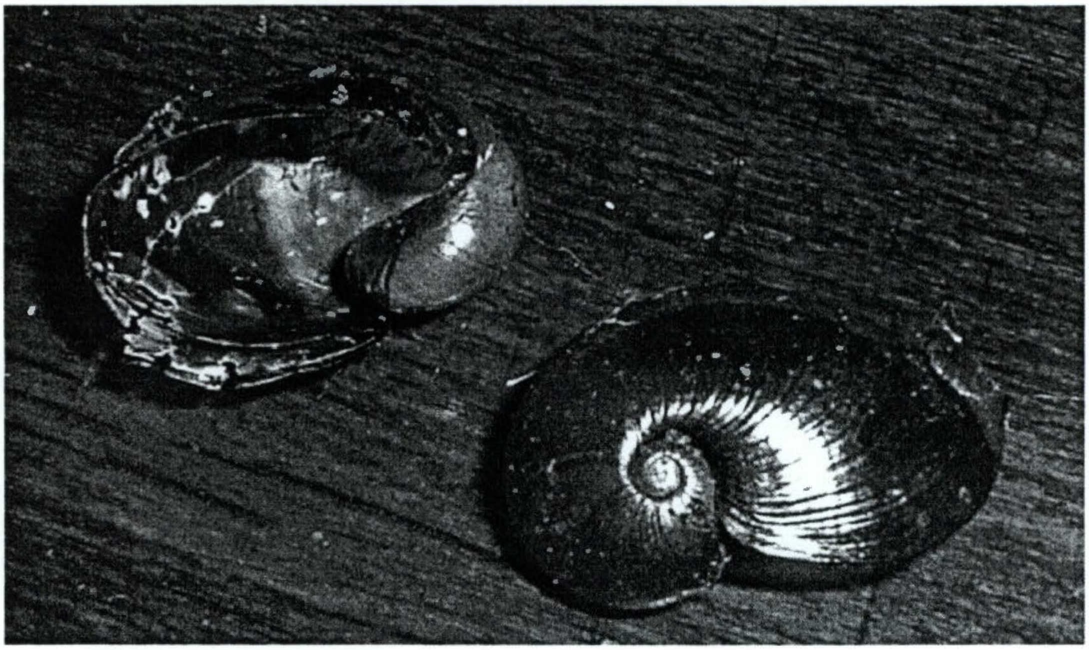


Fig. A.1. Shell of *V. latissimus*

Fig. A.2. *V. latissimus* in GSMNP

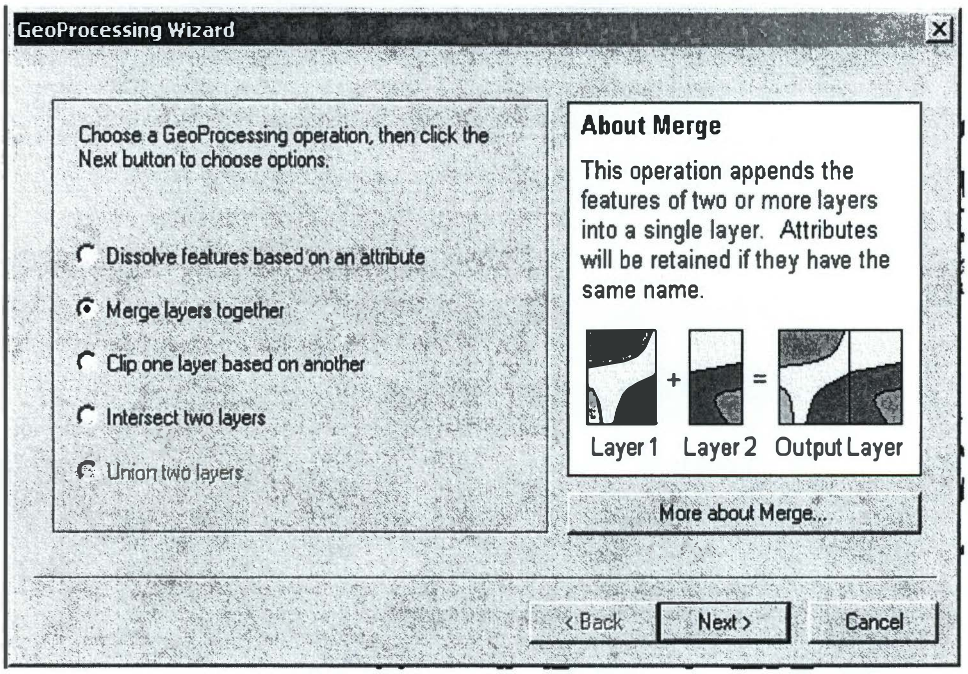
Appendix B
Habitat Suitability Model Tutorial

Habitat Suitability Model Tutorial

B.1 Merging Individual Trail Files into One Master Trail File

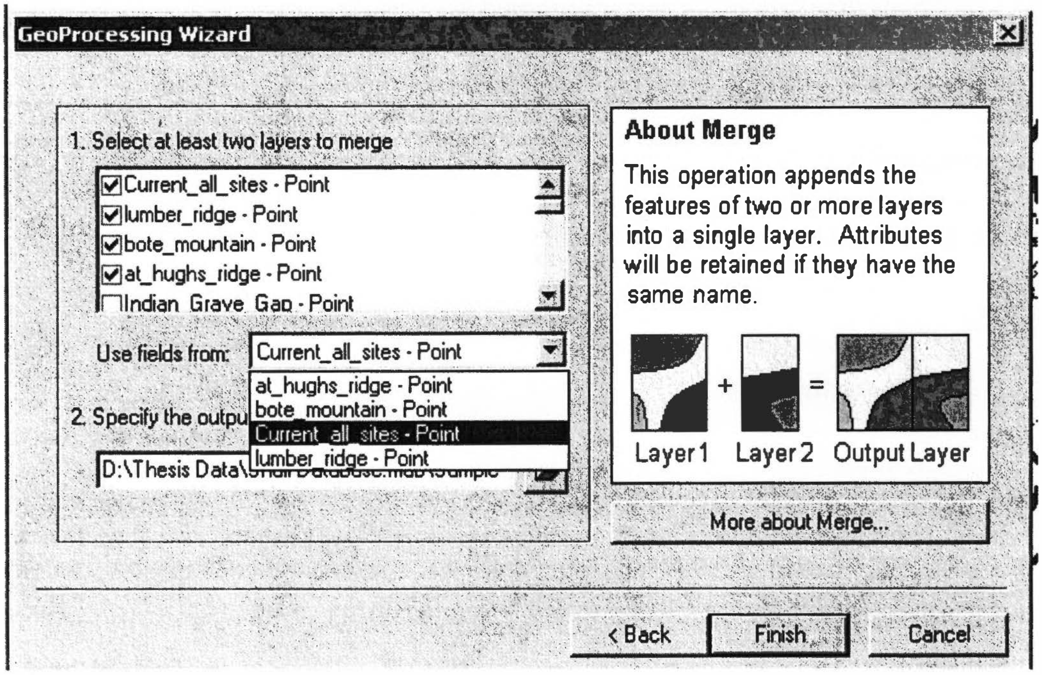
The merge command is used when two or more layers need to be combined into one large layer that contains all the combined features and attributes. Layers that are to be merged must be the same feature type. In this case, we are merging all point features. This process will combine all the individual surveyed trail features into one master trail feature that contains all attribute information collected for each site. To merge the trail features:

1. Under the "Tools" menu, select "GeoProcessing Wizard."



2. Click on the radio button next to "Merge layers together."
3. Click "Next."

4. Select at least two input layers to merge together by checking the box next to the feature name. In this case, choose all the separate trail files that are to be merged into one trail file.



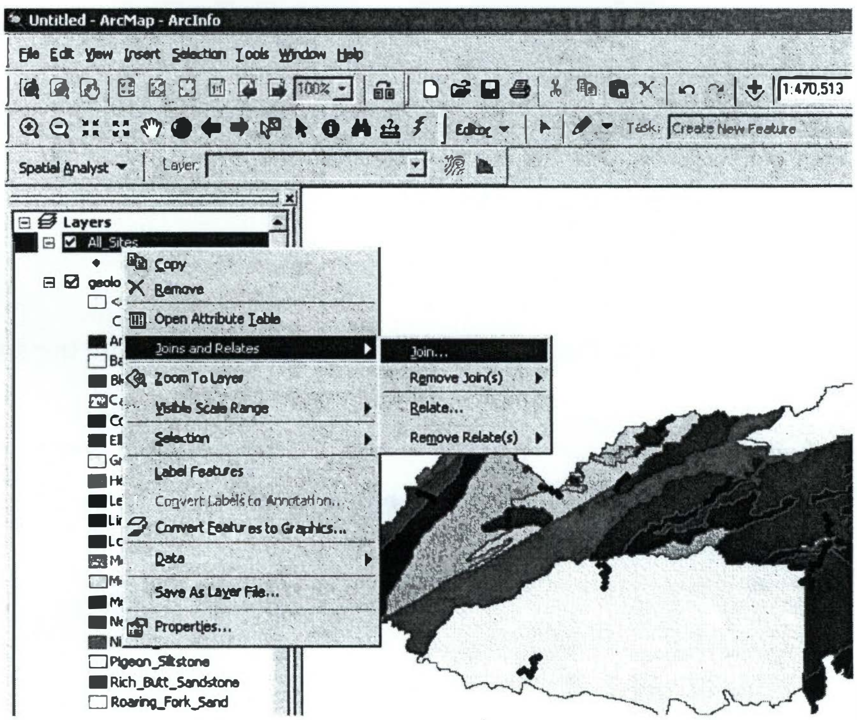
5. Select a layer from which to use the attribute fields from. If a master trails file has already been created, use it for the field form.
6. Specify the output path which should be in the Geodatabase for the project. Creating the merged layer in a predefined feature dataset allows the coordinate system for the merged layer to be automatically defined.
7. Click "Finish" and the Geoprocessing operation will merge the separate trail files into one master trail file.

8. The merged layer will show up in the layout table of contents. Rename the layer and check to make sure all trails and all attributes were merged together.

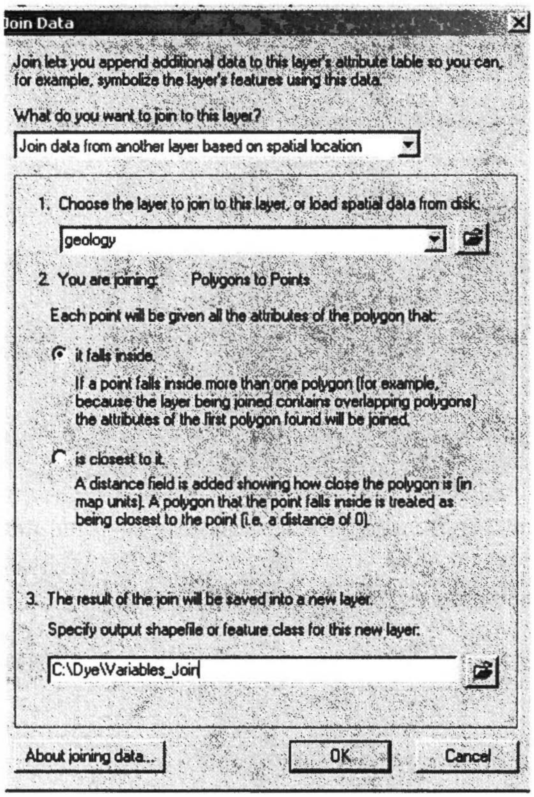
B.2 Spatially Joining Polygon Data to Points

The "Join" command is a very useful tool within ArcGIS. This command attaches attribute information about one variable to the attribute table of another. It is very helpful for combining data to use in a statistical package such as SPSS. For this project, environmental data (including geology, vegetation, soils, and logging history) were spatially joined to the sample site point file. The data table with the combined attributes was then used to conduct statistical analysis. To spatially join polygon data to a point file:

1. Right click on the data layer you wish to join variables to.
2. Click on "Joins and Relates" and choose "Join..."



3. Click the first drop down menu and choose "Join data from another layer based on spatial location."
4. Choose the environmental variable layer to join to the point file.

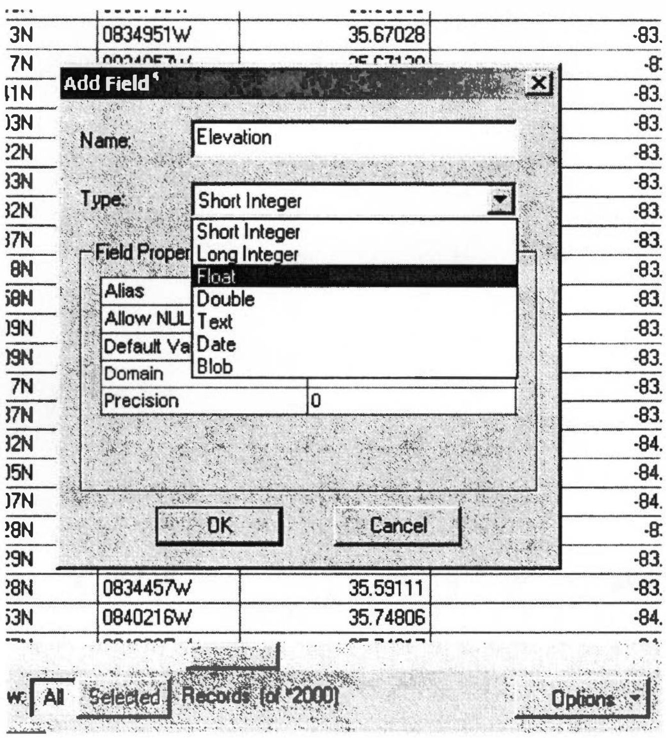


5. Click the first radio button "it falls inside."
6. Specify the output name and location.
7. Click "OK."
8. Open up the attribute layer of the joined feature to make sure the attributes were joined to the table.
9. Repeat for each environmental variable polygon.

B.3 Joining Raster Data to the Point File

It is not possible to convert certain raster data (such as slope, aspect, and elevation) because of the continuous nature of the data. There are too many values to be represented in a large scale vector format. For this study, the continuous raster layers mentioned above were added to the sample site data by hand. To add attribute by hand:

1. In ArcCatalog, create a new field in the attribute table for each raster layer you are adding to the table.
2. Click on the feature class in ArcCatalog and click the "Preview" tab (make sure that "Table" is the active preview option).
3. Under "Options" click "Add Field."
4. Enter a name for the field and make the type "Float" and take the default field properties.



5. In ArcMap, add all raster layers to be added to the feature table into the data frame.

6. Use the "Zoom in" tool  to view a single set of sample sites along a trail.

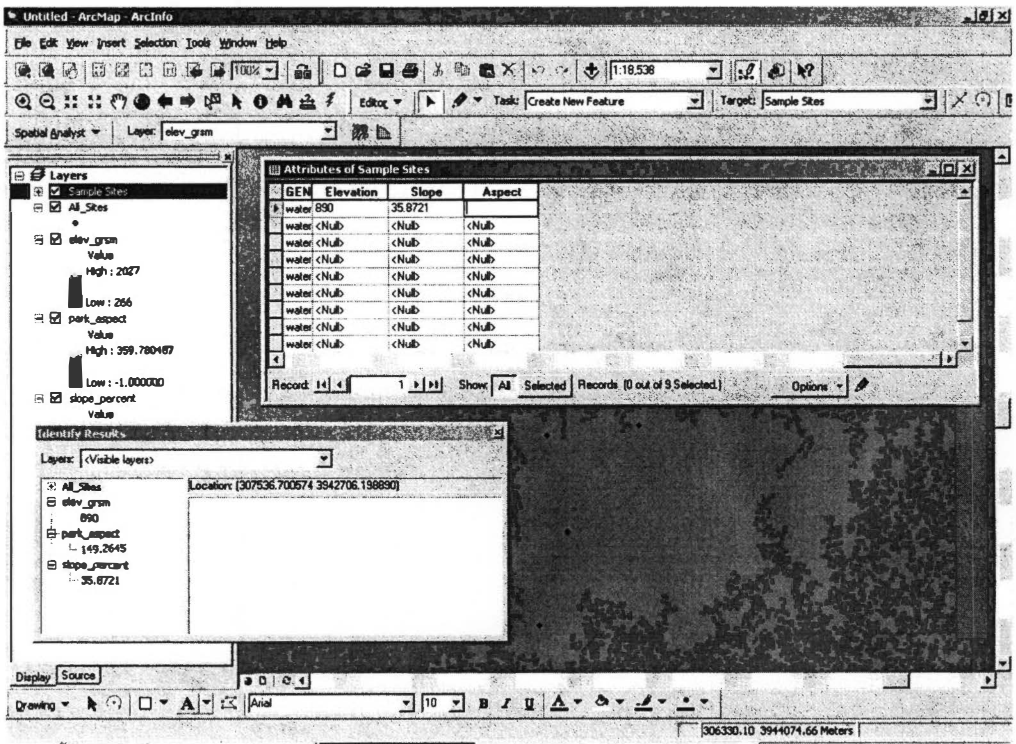
7. Under Editor click "Start Editing."

8. Make the "Identify" tool active .

9. Open the attribute table of the layer to which you are adding data.

10. Click on a sample site and the "Identify Results" window will appear.

11. Under the drop down menu, click "Visible Layers" and all values will appear.



12. Enter the elevation, slope, and aspect values under the respective fields for each sample site.

13. Repeat this procedure for all trails.

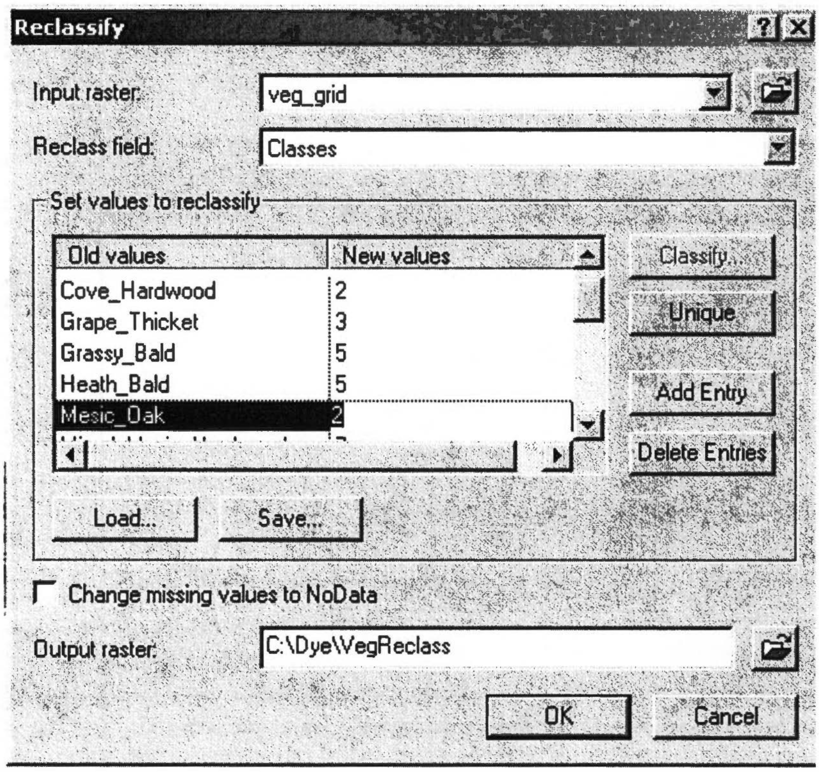
14. When all trails have the added raster attributes, click "Save Edits" in Editor.

B.4 Reclassifying Environmental Factor Grids

To analyze the habitat variables within the habitat suitability model, the variables must be reclassified based on their relationships with the presence of the species. Rankings are based on statistical analysis, and each variable should have the same number of rankings. In this study, the three significant variables (elevation, soils, and vegetation) were reclassified into five rankings. A ranking of one represents classes that are least suitable for the species and a ranking of five represents the optimal classes for the species. To reclassify significant habitat variables:

1. Rank the classes within each variable according to the statistical relationships with the presence of the species. One should be the lowest score and five should be the highest.
2. Add all significant variables to ArcMap in raster format.
3. Turn on the Spatial Analyst extension by clicking on "Tools" and clicking "Customize." Click the box next to Spatial Analyst to make it visible. Click "Tools" and click "Extensions" to make sure that the extension is turned on.
4. Click on the "Spatial Analyst" menu and click on "Reclassify…"

5. Make sure that the input raster and Reclass fields are correct for the layer you are reclassifying.

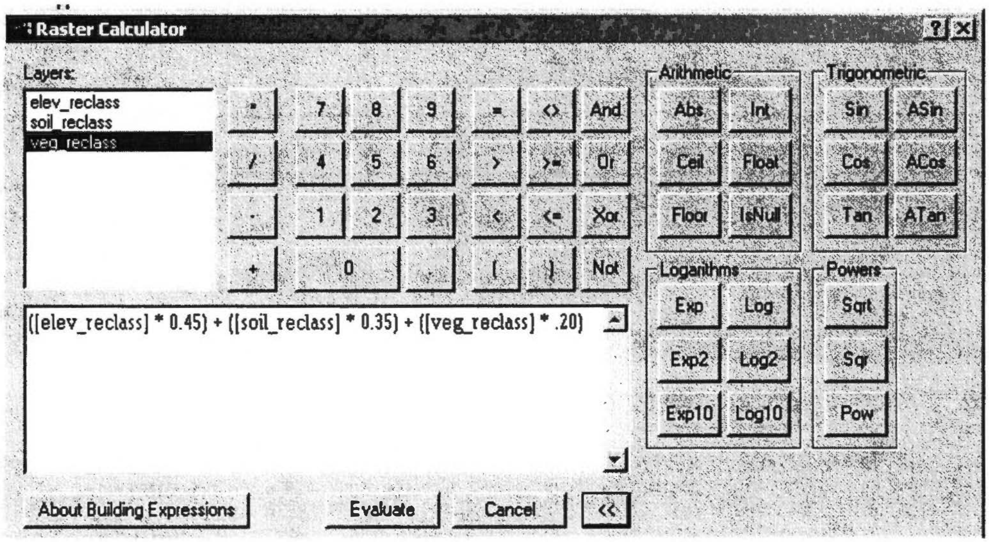


6. Based on your rankings of each class, input the suitability rankings into the "New Values" field.
7. Specify the output raster name and location.
8. Click "OK."
9. The reclassified layer will appear in the data view within ArcMap. Make sure that the new values are correct.
10. Reclassify each significant variable in the same manner.

B.5 Evaluating Habitat Suitability Areas

Once the data have been reclassified, the data are ready to be input into the raster calculator to produce the habitat suitability model. Weights can be assigned to variables based on the strength of relationship between the variable and the presence of the organism. In this study, weights were assigned to reclassified elevation, vegetation, and soil grids. Weights were based on logistic regression results conducted on the data. To make a habitat suitability model:

1. Add all significant reclassified grid layers to the data frame in ArcMap.
2. Turn on the Spatial Analyst extension by clicking on "Tools" and clicking "Customize." Click the box next to Spatial Analyst to make it visible. Click "Tools" and click "Extensions" to make sure that the extension is turned on.
3. Under Spatial Analyst click on "Raster Calculator."
4. Enter the expression to build the model in the dialog box in "Raster Calculator."
    A. Open brackets and double click on the first reclassed variable.
    B. Multiply the variable by the selected weight and close the brackets.
    C. Click the plus (+) sign.
    D. Open a bracket and double click the next variable.
    E. Multiply the variable by the selected weight and close brackets.
    F. Repeat for all variables included in the model.



5. Once the expression is built, click the "Evaluate" button and the calculation will begin.
6. Once the calculation is finished, the habitat suitability map will be output in the data frame in ArcMap.
7. Once the model has been made, analysis can be conducted on the habitat areas through a variety of GIS-based methods.

Vita

Andrew Strom Dye was born in Nashville, Tennessee on April 28, 1978. Growing up he enjoyed countless trips to the Great Smoky Mountains with his family and friends, which instilled a great love for the park and its wondrous diversity. He graduated from Brentwood Academy in the spring of 1997 and went on to spend four great years in Birmingham, Alabama attending Samford University as a student and a track and field athlete. During a memorable college experience he had opportunities to study in Western Europe and the Peruvian Amazon before he graduated from Samford in the spring of 2001 with a Bachelor of Science degree in Environmental GIS. In August 2001, Andrew moved to Knoxville and started his pursuit of a Master of Science degree in Geography at The University of Tennessee. In Knoxville he worked as a consultant for the Tennessee Geographic Alliance, as a Teaching Assistant for Introduction to World Geography, and as a Graduate Assistant in the President's Office. During his time in Knoxville, he has had the chance to travel through the British Isles, Kenya and Tanzania, the eastern Caribbean, and is currently preparing to depart for New Zealand and Australia. He leaves with the graduating class of 2004 ready to begin the next chapter of his life.

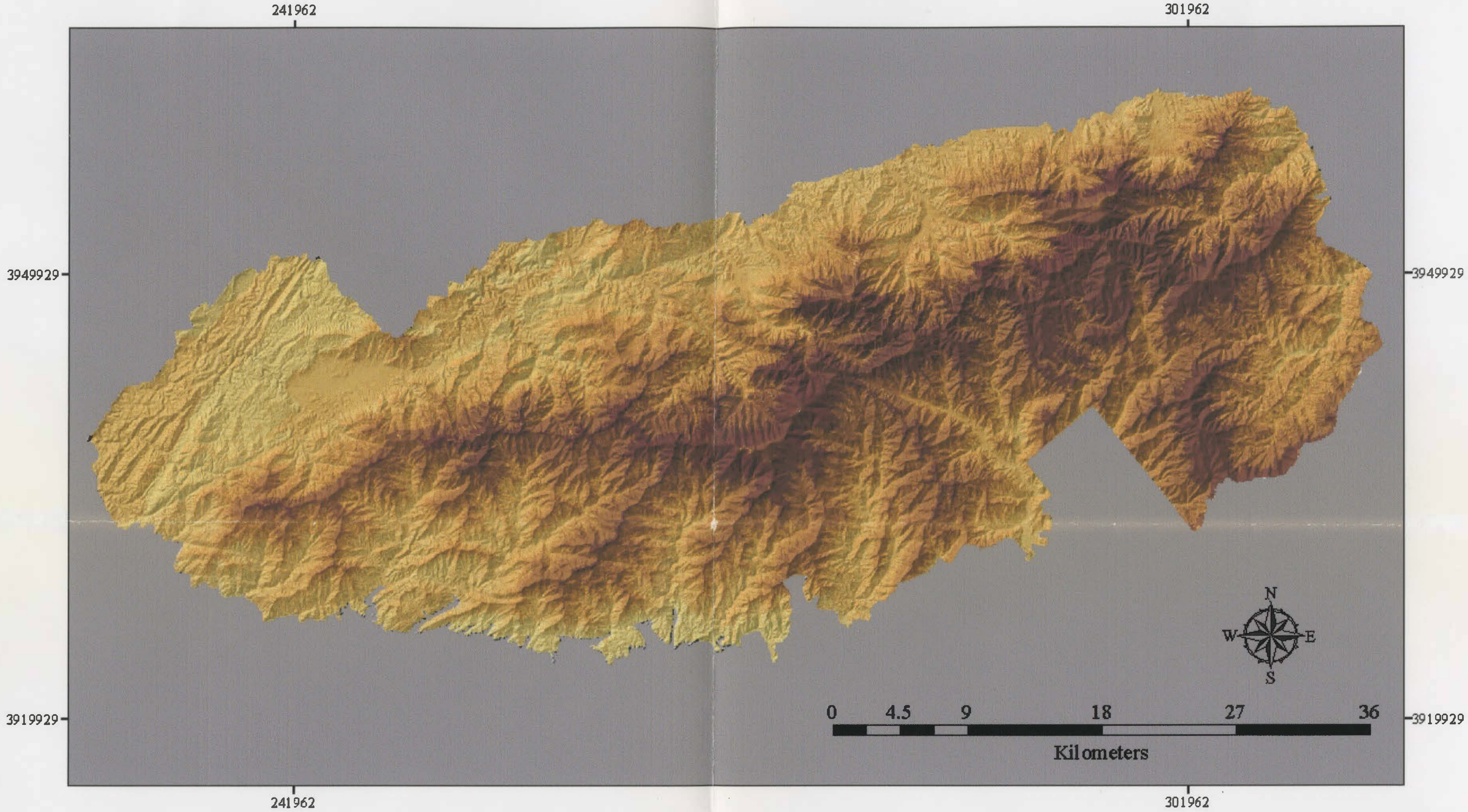


Plate 1. Habitat Suitability Model for V. latissimus in Great Smoky Mountains National Park

Andrew Strom Dye

August 2004

Habitat Suitability

Highest Ranked Areas

Lowest Ranked Areas